The Dissertation Committee for Tianyang Bai
certifies that this is the approved version of the following dissertation:

Analysis of Millimeter Wave and Massive MIMO Cellular Networks

Committee:

Robert W. Heath, Jr., Supervisor

Jeffrey G. Andrews

François Baccelli

Lili Qiu

Sujay Sanghavi

Analysis of Millimeter Wave and Massive MIMO Cellular Networks

by

Tianyang Bai, B.E., M.S.E.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2016

Dedicated to my parents,

Hong and Shizhen.

Acknowledgments

I can only start with my deepest gratitude to my Ph.D. supervisor Prof. Robert W. Heath, Jr. for offering me an opportunity five years ago, and having faith in me ever since. Without his support and technical guidance, I would have never completed my dissertation. While all Ph.D. students may have their own definitions of the best supervisor, I consider myself the luckiest, as I have worked with my own for five years. I would like to thank Dr. Heath for tolerating my shortcomings but never losing patience in helping me overcome them. In addition, I always regard Dr. Heath as the role model of my career for his great passion and technical excellence. Having a high self-expectation, I would still consider my career a great success if I could accomplish half of his achievements so far.

I would like to express my sincere gratitude to Prof. François Baccelli for his great series of lectures on stochastic geometry as well as his wise advice and comments on my research. Many of my research ideas originated from what I learned in his classes. He also taught me the beauty of mathematics in generality. I also owe a great debt of gratitude to Prof. Jeffrey Andrews for his great pioneering papers on stochastic geometry that I have closely followed during my PhD study. I would like to thank Prof. Lili Qiu and Prof. Sujay Sanghavi for their intellectually stimulating classes and their comments on my

dissertation.

It was a great pleasure to work with my gifted labmates and visitors at the WSIL: Omar El Ayach, Salam Akoum, Kien Trung Truong, Amin Abdel Khalek, Nachiappan Valliappan, Chao Chen, Rahul Vaze, Ken Pesyna, Andrew Thornburg, Preeti Kumari, Talha Ahmed Khan, Jianhua Mo, Vutha Va, Kiran Venugopal, Yueping Wu, Marios Kountouris, Martin Taranetz, Xiao Li, and Christopher Mollen. A special mention is due to Namyoon Li, as I always learnt from chatting with him both in research and life. I would like to thank the postdoc fellows Xinchen Zhang and Junil Choi for their valuable comments and discussions on my research.

I am also very lucky to meet with my WNCG friends: Yuhuan Du, Yicong Wang, Zheng Lu, Jiaoxiao Zheng, Hongbo Si, Qiaoyang Ye, Sarabjot Singh, Xingqin Lin, Yingzhe Li, and Mandar Kulkarni, who have made all my days at UT much fun. As all the WNCG students, I want to thank Melanie Gulick and Karen Little for their help in the administrative work.

I would like to thank Dr. Sundar Subramanian for the opportunity to work with the talented mmWave team at Qualcomm during my internships. I really enjoyed the productive and fun summers in New Jersey.

Last, but definitely not least, I am forever grateful to my family for their unconditional support and constant encouragement. I dedicate everything that I have accomplished to them. Words fail me in expressing my love to them .

Analysis of Millimeter Wave and Massive MIMO Cellular Networks

Publication No. _____

Tianyang Bai, Ph.D.
The University of Texas at Austin, 2016

Supervisor: Robert W. Heath, Jr.

Millimeter wave (mmWave) communication and massive multiple-input multiple-output (MIMO) are promising techniques to increase system capacity in 5G cellular networks. The prior frameworks for conventional cellular systems do not directly apply to analyze mmWave or massive MIMO networks, as (i) mmWave cellular networks differ in the different propagation conditions and hardware constraints; and (ii) with a order of magnitude more antennas than conventional multi-user MIMO systems, massive MIMO systems will be operated in time-division duplex (TDD) mode, which renders pilot contamination a primary limiting factor.

In this dissertation, I develop stochastic geometry frameworks to analyze the system-level performance of mmWave, sub-6 GHz massive MIMO, and mmWave massive MIMO cellular networks. The proposed models capture the key features of each technique, and allow for tractable signal-to-interference-plus-noise ratio (SINR) and rate analyses. In the first contribution, I develop

an mmWave cellular network model that incorporates the blockage effect and directional beamforming, and analyze the SINR and rate distributions as functions of the base station density, blockage parameters, and antenna geometry. The analytical results demonstrate that with a sufficiently dense base station deployment, mmWave cellular networks are capable to achieve comparable SINR coverage and much higher rates than conventional networks. In my second contribution, I analyze the uplink SINR and rate in sub-6 GHz massive MIMO networks with the incorporation of pilot contamination and fractional power control. Based on the analysis, I show scaling laws between the number of antennas and scheduled users per cell that maintain the uplink signal-to-interference ratio (SIR) distributions are different for maximum ratio combining (MRC) and zero-forcing (ZF) receivers. In my third contribution, I extend the sub-6 GHz massive MIMO model to mmWave frequencies, by incorporating key mmWave features. I leverage the proposed model to investigate the asymptotic SINR performance, when the number of antennas goes to infinity. Numerical results show that mmWave massive MIMO outperforms its sub-6 GHz counterpart in cell throughput with a dense base station deployment, while the reverse can be true with a low base station density.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview

The coming 5G cellular networks will apply novel communication techniques, including millimeter wave (mmWave) communications and massive multiple-input and multiple-output (MIMO), to increase system capacity [42]. The mmWave frequency refers to the spectrum ranging from 30 GHz to 300 GHz, while conventionally the spectrum above 20 GHz, e.g. the local multipoint distribution service (LMDS) band at 28 GHz, is categorized into mmWave as well. Operating wireless systems at mmWave frequencies is by no means a new topic, as the study of mmWave goes back to 1890s [54]. In consumer radios, the mmWave spectrum has been applied in personal area networking (WirelessHD and IEEE 802.15.3c)[48, 35], local area networking (IEEE 802.11ad) [152], and fixed-point backhaul links in cellular networks [69].

Recently, mmWave has drawn great interest as a new carrier frequency for access channels in 5G cellular networks [112, 120, 17, 42]. Compared with the bandwidth-limited conventional spectrum below 6 GHz, mmWave can potentially provide several gigahertz bandwidths to solve the spectrum gridlock in cellular networks [112, 120]. For example, the Federal Communications

Commission (FCC) in the USA has been considering making rules to authorize mobile operations in the LMDS band at 27.5-28.35 GHz and the 39 GHz band at 38.6- 40 GHz with county-size licenses [47]. Besides, the unlicensed band at 57- 64 GHz and E-band at 71-76 GHz and 81-86 GHz are also likely to be used for future cellular networks, due to the large continuous bandwidth [112, 120, 92].

MmWave cellular networks will operate in a different manner from conventional cellular systems below 6 GHz. To begin with, measurements reveal different channel statistics at mmWave from those at sub-6 GHz frequencies [120, 121]: the path loss between single dipole antennas increases quadratically with the carrier frequency, due to the smaller aperture size; high sensitivity to blockages is found at mmWave, due to high penetration loss through certain materials and weak diffraction [113, 120]. Moreover, mmWave cellular networks will apply different transceiver architectures for signal processing, due to the hardware and power constraints [122, 10]. For one thing, large antenna arrays will be deployed for direction beamforming at both base stations and mobile stations, to increase the effective aperture size and meet the link budget [112, 42, 122]. For another, with large antenna arrays and wide bandwidths, more energy-efficient transceiver architectures will be applied in mmWave, to meet the constraints from the circuity power consumption [122]. For instance, mmWave cellular networks may apply analog or hybrid beamforming to reduce the number of RF chains, and use low-resolution ADCs at receivers to reduce the sampling power [122, 10].

Massive MIMO is another promising physical layer technology to increase cell throughput in fifth generation cellular systems [90, 76, 87, 124, 42]. Compared with the conventional MU-MIMO systems, massive MIMO systems deploy a order of magnitude more antennas, e.g. 64 or more antennas, at base stations to simultaneously serve a dozen of users in their cells [90]. Besides the gain in cell throughout, theoretical analyses have also shown the following advantages when the number of antennas goes large: (a) the effects of small-scaling fading vanish with infinite antennas, as the SINR converges to a deterministic equivalence [90, 64]; the impacts of interference and noise become minor, as under certain assumptions, e.g. IID Rayleigh fading channel, the desired signal channel becomes asymptotically orthogonal to noise and inference channels [90, 64, 105]; simple linear signal processing provides near-optimal performance with large antenna arrays [124]; compared with the single-input single-output (SISO) case, less transmit power is needed to achieve the same spectrum efficiency, as it scales down with the number of antennas [104].

The exact spectrum for deploying massive MIMO, however, is still not clear: the conventional spectrum below 6 GHz, e.g. the ultra high frequency (UHF) frequency, has been the defacto operation band for massive MIMO, but has a limited bandwidth; the mmWave band may potentially offer a larger bandwidth, but the performance will differ from the conventional massive MIMO, due to different propagation conditions and hardware designs [120].

In this dissertation, I propose to analyze the SINR coverage and rate

in mmWave and massive MIMO cellular networks using the tool of stochastic geometry. In the rest of this chapter, before presenting the main technical results, I will introduce basic concepts and research surveys for massive MIMO and mmWave communications in Section 1.2- Section1.5; then I summarize the contributions in Section 1.6. In Chapter 2, I introduce the system model and key results on analyzing downlink mmWave cellular networks. The analysis of uplink massive MIMO networks in the sub-6 GHz bands is provided in Chapter 3. Then I extend the framework in Chapter 3 to analyze mmWave massive MIMO cellular networks in Chapter 4. Finally, I conclude the dissertation, and discuss directions for future work in Chapter 5.

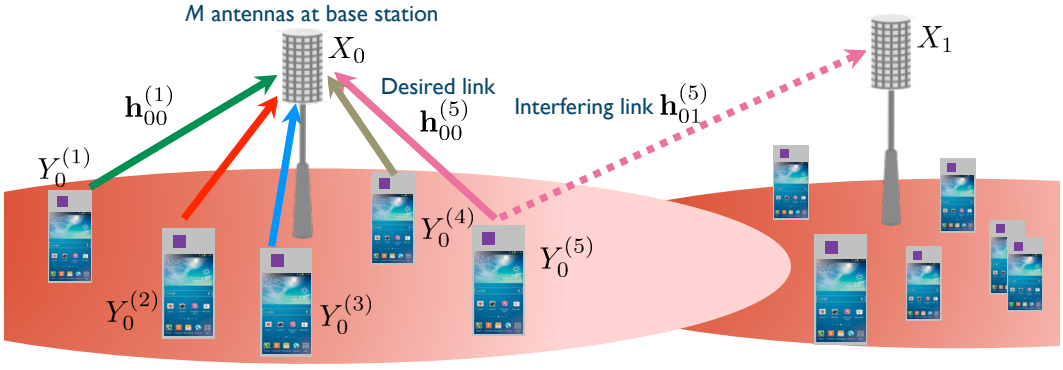


Figure 1.1: Massive MIMO network with $L = 2$ cells and $K = 5$ users per cell. I denote the location of the ℓ-th base station as X_ℓ, k-th user in the ℓ-th cell as $Y_\ell^{(k)}$, and the channel vector from $Y_\ell^{(k)}$ to X_ℓ' as $\mathbf{h}_{\ell'\ell}^{(k)}$.

1.2 The Massive MIMO Concept

In this section, I first illustrate the concept of massive MIMO systems using a basic example system in Section 1.2.1. Then, I introduce three operation stages of a time-division duplex (TDD) massive MIMO networks: uplink channel training, uplink data transmission, and downlink data transmission in Section 1.2.2-1.2.4. Finally, in Section 1.2.5, I introduce initial results on the asymptotic performance of massive MIMO, when the number of antennas goes to infinity, based on [90].

1.2.1 Massive MIMO System

Now I consider a MU-MIMO cellular system consisting of L cells, where in each cell, the base station has M antennas to serve K single-antenna mobile stations in a resource block. In a massive MIMO system, I generally assume $M \gg K$, e.g., $M = 128$ antennas serve $K = 10$ users in a cell. I show an example of $L = 2$ and $K = 5$ as an example in Fig. 1.1. For $\ell \in [0, L-1]$, and $k \in [1, K]$, let X_ℓ be the location of the ℓ-th base station, and $Y_\ell^{(k)}$ be the location of the k-th scheduled user in the ℓ-th cell. To model the prorogation channel, I use a M-dimension vector $\mathbf{h}_{\ell'\ell}^{(k)}$ to represent the channel from $Y_\ell^{(k)}$ to $X_{\ell'}$. In this section, to illustrate the basic concepts of massive MIMO systems, I assume the *IID Rayleigh fading channel model*: the channel vector $\mathbf{h}_{\ell'\ell}^{(k)}$ can be computed as

$$\mathbf{h}_{\ell'\ell}^{(k)} = \beta_{\ell'\ell}^{(k)} \mathbf{w}_{\ell'\ell}^{(k)}, \tag{1.1}$$

where $\beta_{\ell'\ell}^{(k)}$ is a complex constant accoutring for effects including large-scale path loss and shadowing, $\mathbf{w}_{\ell'\ell}^{(k)}$ is an M-dimensional random Gaussian vector of the distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I_M})$, and $\mathbf{I_M}$ represents an M-dimensional identity matrix. For ease of notation, I define an channel matrix $\mathbf{H}_{\ell'\ell} \in \mathcal{C}^{M\times K}$ as

$$\mathbf{H}_{\ell'\ell} = \left[\mathbf{h}_{\ell'\ell}^{(1)},\ ...,\ \mathbf{h}_{\ell'\ell}^{(K)}\right], \tag{1.2}$$

where its k-th column is the channel vector $\mathbf{h}_{\ell'\ell}^{(k)}$. Note that $\mathbf{H}_{\ell'\ell}$ represents the uplink channels from all users in ℓ'-th cell to base station X_ℓ.

In order to design the combining and precoding vectors for the uplink and downlink data transmission, the channel state information (CSI) of the desired signal link is obtained through channel training. In this monograph, I focus on the massive MIMO systems operated in the *TDD mode* [90], where the reciprocity between the uplink and downlink channels is exploited to reduce the training overhead. In a TDD system, the amount of channel training overhead is proportional to the number of scheduled users K, while in a frequency-division duplex (FDD) system, if no special structure of the channel covariance is assumed, the downlink training overhead is in general proportional to the number of the base station antennas M [90]. Moreover, in a FDD system, it requires additional feedback of the downlink CSI from users to their base stations, as the downlink channel can only be estimated at mobile stations. The complexity and overhead of the CSI feedback can also be high, when M grows large. Typically, a TDD massive MIMO system is operated in three stages: uplink channel training, downlink data transmission, and uplink data transmission [90], as described in the following sections.

1.2.2 Uplink channel training

In a TDD massive MIMO system, channel training is performed in the uplink: pilot sequences are transmitted by scheduled users to base stations to estimate channels [90]. Here I consider a simple case as in [90]: all scheduled users in a cell are assigned with orthogonal pilots $\mathbf{t}^{(k)}$ of length τ; the same set of pilot sequences $\mathbf{T}$ is reused in every cell, where the matrix

$$\mathbf{T} = \left[\mathbf{t}^{(1)},\ ...,\ \mathbf{t}^{(K)}\right], \tag{1.3}$$

and satisfies $\mathbf{T}^{\mathrm{H}}\mathbf{T} = \tau\mathbf{I}_K$; perfect synchronization of pilot transmission is assumed across the network. More complicated designs to reduce the estimation error are discussed in Section 1.3.2.

In the channel training stage, all scheduled users transmit their assigned pilots, and base station X_ℓ receives

$$\mathbf{Y}_\ell = \sqrt{P_{\mathrm{T}}}\sum_{\ell'=1}^{L}\mathbf{H}_{\ell\ell'^{\mathrm{T}}}\mathbf{T}' + \mathbf{N}_\ell, \tag{1.4}$$

where P_{T} is the transmit power for channel training, $\mathbf{N}_\ell$ represents the thermal noise at base station X_ℓ, and each entry of $\mathbf{N}_\ell$ follows the distribution $\mathcal{CN}(0, \sigma^2)$.

Next, base station X_ℓ correlates the received signal $\mathbf{Y}_\ell$ with the pilots as

$$\mathbf{U}_\ell = \frac{1}{\tau}\mathbf{Y}_\ell\mathbf{T}' \tag{1.5}$$

$$= \sum_{\ell'=1}^{L}\sqrt{P_{\mathrm{T}}}\mathbf{H}_{\ell\ell'} + \frac{1}{\tau}\mathbf{N}_\ell\mathbf{T}' \tag{1.6}$$

$$= \sum_{\ell'=1}^{L} \sqrt{P_{\mathrm{T}}} \mathbf{H}_{\ell\ell'} + \bar{\mathbf{N}}_{\ell}, \tag{1.7}$$

where $\bar{\mathbf{N}}_{\ell} = \frac{1}{\tau}\mathbf{N}_{\ell}\mathbf{T}'$, and each entry of $\bar{\mathbf{N}}_{\ell}$ follows the distribution $\mathcal{CN}(0, \frac{\sigma^2}{\tau})$.

For ease of illustration, denote the k-th column of $\mathbf{U}_{\ell}$ as $\mathbf{u}_{\ell}^{(k)}$, and it follows that

$$\mathbf{u}_{\ell}^{(k)} = \sqrt{P_{\mathrm{T}}}\mathbf{h}_{\ell\ell}^{(k)} + \sum_{\ell' \neq \ell}^{L} \sqrt{P_{\mathrm{T}}}\mathbf{h}_{\ell\ell'}^{(k)} + \bar{\mathbf{n}}_{\ell}^{(k)}. \tag{1.8}$$

The vector $\mathbf{u}_{\ell}^{(k)}$ can be viewed as the observation of the channel $\mathbf{h}_{\ell\ell}^{(k)}$ at base station X_{ℓ}. Note that besides the contribution from the desired signal channel $\mathbf{h}_{\ell\ell}^{(k)}$, the observation $\mathbf{u}_{\ell}^{(k)}$ also includes the thermal noise term, and more importantly, the interference term $\sum_{\ell' \neq \ell}^{L} \sqrt{P_{\mathrm{T}}}\mathbf{h}_{\ell\ell'}^{(k)}$ coming from other-cell users using the same pilot $\mathbf{t}^{(k)}$.

Next, I assume the large-scale path loss information $\beta_{\ell\ell'}^{(k)}$ is known at base stations. Then base station X_{ℓ} applies an MMSE estimator to estimate the channel $\mathbf{h}_{\ell\ell}^{(k)}$ as

$$\bar{\mathbf{h}}_{\ell\ell}^{(k)} = P_{\mathrm{t}}\beta_{\ell\ell}^{(k)}\mathbf{I}_M \left[\frac{\sigma^2}{\tau}\mathbf{I}_M + \left(\sum_{\ell'=1}^{L} P_{\mathrm{t}}\beta_{\ell\ell'}^{(k)}\right)\mathbf{I}_M\right]^{-1} \mathbf{u}_{\ell}^{(k)} \tag{1.9}$$

$$= \frac{\beta_{\ell\ell}^{(k)}}{\frac{\sigma^2}{P_{\mathrm{t}}\tau} + \sum_{\ell'=1}^{L}\beta_{\ell\ell'}^{(k)}}\mathbf{u}_{\ell}^{(k)}, \tag{1.10}$$

where $\bar{\mathbf{h}}_{\ell\ell}^{(k)}$ is the estimate of $\mathbf{h}_{\ell\ell}^{(k)}$; By the orthogonal principle of the MMSE estimation, the estimation error $\hat{\mathbf{h}}_{\ell\ell}^{(k)} = \mathbf{h}_{\ell\ell}^{(k)} - \bar{\mathbf{h}}_{\ell\ell}^{(k)}$ is uncorrelated with $\hat{\mathbf{h}}_{\ell\ell}^{(k)}$, and follows the distribution $\mathcal{CN}\left(\mathbf{0}, \beta_{\ell\ell}^{(k)}\left(1 - \frac{\beta_{\ell\ell}^{(k)}}{\frac{\sigma^2}{\tau P_{\mathrm{t}}} + \sum_{\ell'=1}^{L}\beta_{\ell\ell'}^{(k)}}\right)\mathbf{I}_M\right)$.

Note that in the IID Rayleigh fading case, the channel estimate $\bar{\mathbf{h}}_{\ell\ell}^{(k)}$ is a scaled version of the vector $\mathbf{u}_{\ell}^{(k)}$. Therefore, by (1.8), even if ignoring the thermal noise, the channel estimator still contains errors due to the interference from the co-pilot users in the other cells— a phenomenon that I call *pilot contamination*. In TDD massive MIMO systems, pilot contamination is a limiting factor for the asymptotic SINR and rate performance, when the number of antennas goes to infinity [90]. I are going to discuss the techniques to reduce or mitigate the impact of pilot contamination in Section 1.3.2.

1.2.3 Uplink Data Transmission

Now I continue to investigate the uplink SINR and ergodic rate performance. In the uplink data transmission, the received signal at base station X_ℓ is

$$\mathbf{y}_\ell = \sqrt{P_{\mathrm{u}}} \sum_{\ell',n} \mathbf{h}_{\ell\ell'}^{(n)} s_{\ell'}^{(n)} + \mathbf{n}_\ell, \tag{1.11}$$

where P_{u} is the uplink transmit power; $s_{\ell'}^{(n)}$ is the symbol from user $X_{\ell}^{(n)}$, and satisfies the unit power constraint $\mathbb{E}\left[|s_{\ell'}^{(n)}|^2\right] = 0$; $\mathbf{n}_\ell$ is the noise vector following the distribution $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$.

To illustrate the basic concepts, in this section, I focus on the performance of MRC receivers, while I summarize the results for more complicated receivers in Section 1.3.1. With MRC receivers, base station X_ℓ treats the channel estimate $\bar{\mathbf{h}}_{\ell\ell}^{(k)}$ as the true channel, and uses the combining vector

$\mathbf{g}_\ell^{(k)} = \bar{\mathbf{h}}_{\ell\ell}^{(k)}$ to decode symbol $s_\ell^{(k)}$ as

$$\bar{s}_\ell^{(k)} = \underbrace{\sqrt{P_{\mathrm{u}}}\mathbf{g}_\ell^{(k)*}\bar{\mathbf{h}}_{\ell\ell}^{(k)}s_\ell^{(k)}}_{\text{signal}} + \underbrace{\sqrt{P_{\mathrm{u}}}\mathbf{g}_\ell^{(k)*}\hat{\mathbf{h}}_{\ell\ell}^{(k)}s_\ell^{(k)}}_{\text{channel estimation error}} + \underbrace{\sqrt{P_{\mathrm{u}}}\sum_{\ell'\neq\ell,n\neq k}\mathbf{g}_\ell^{(k)*}\mathbf{h}_{\ell\ell'}^{(n)}s_{\ell'}^{(n)}}_{\text{interference}} + \underbrace{\mathbf{g}_\ell^{(k)*}\mathbf{n}_\ell}_{\text{noise}}. \quad (1.12)$$

Note that the terms due to channel estimation error, interference, and noise are unknown to base station X_ℓ. Therefore, the uplink SINR for user $Y_\ell^{(k)}$ is expressed as

$$\mathrm{SINR}_{\mathrm{u}} = \frac{|\mathbf{g}_\ell^{(k)*}\bar{\mathbf{h}}_{\ell\ell}^{(k)}|^2}{\mathbb{E}|\mathbf{g}_\ell^{(k)*}\hat{\mathbf{h}}_{\ell\ell}^{(k)}|^2 + \sum_{\ell'\neq\ell,n\neq k}|\mathbf{g}_\ell^{(k)*}\mathbf{h}_{\ell\ell'}^{(n)}|^2 + |\mathbf{g}_\ell^{(k)}|^2\frac{\sigma^2}{P_{\mathrm{u}}}}, \quad (1.13)$$

where the second term in the denominator is due to the error of the MMSE channel estimator, and the expectation is taken with respect to the error $\tilde{\mathbf{h}}_{\ell\ell}^{(k)}$. Further, by treating the interference and noise terms in (1.13) as additive Gaussian noise independent of $s_\ell^{(k)}$, I can obtain a lower bound on the achievable ergodic rate [104, 70] as

$$\xi_{\mathrm{u}} = \mathbb{E}[\log_2(1 + \mathrm{SINR}_{\mathrm{u}})], \quad (1.14)$$

where the expectation is taken with respect to different realizations of channels and noise.

1.2.4 Downlink Data Transmission

To transmit downlink data, base stations apply certain precoding based on the channel estimate. In this section, I focus on the match-filtering precoding: the precoder for the k-th user in the ℓ-th cell is $\mathbf{f}_\ell^{(k)} = \frac{\hat{\mathbf{h}}_\ell^{(k)}}{|\hat{\mathbf{h}}_\ell^{(k)}|^2}$. Since the

channel estimation is performed in the uplink, I assume the user $X_\ell^{(k)}$ has no instantaneous channel state information, but only know the average effective channel gain $\mathbb{E}\left[\mathbf{h}_{\ell\ell}^{(k)*}\mathbf{f}_\ell^{(k)}\right]$. Let $s_\ell^{(k)}$ be the symbol for user $X_\ell^{(k)}$. Then the received signal at user $X_\ell^{(k)}$ is

$$\bar{s}_\ell^{(k)} = \sqrt{P_{\mathrm{d}}}\mathbb{E}\left[\mathbf{h}_{\ell\ell}^{(k)*}\mathbf{f}_\ell^{(k)}\right] s_\ell^{(k)} + \underbrace{\sqrt{P_{\mathrm{d}}}\left(\mathbf{h}_{\ell\ell}^{(k)*}\mathbf{f}_\ell^{(k)} - \mathbb{E}\left[\mathbf{h}_{\ell\ell}^{(k)*}\mathbf{f}_\ell^{(k)}\right]\right) s_\ell^{(k)} + \sqrt{P_{\mathrm{d}}}\sum_{(n,\ell')\neq(k,\ell)} \mathbf{h}_{\ell\ell'}^{(n)*}\mathbf{f}_{\ell'}^{(n)} s_{\ell'}^{(n)} + \mathbf{n}_{\mathrm{d}}}_{\text{unknown to user } Y_\ell^{(k)}}, \tag{1.15}$$

where $\mathbf{n}_{\mathrm{d}}$ is the noise vector for downlink transmission. Then the downlink SINR for user $Y_\ell^{(k)}$ can be expressed as

$$\mathrm{SINR}_{\mathrm{d}} = \frac{\left|\mathbb{E}\left[\mathbf{h}_{\ell\ell}^{(k)*}\mathbf{f}_\ell^{(k)}\right]\right|^2}{\mathbf{Var}\left[\mathbf{h}_{\ell\ell}^{(k)*}\mathbf{f}_\ell^{(k)}\right] + \sum_{(n,\ell')\neq(k,\ell)}\left|\mathbf{h}_{\ell'\ell}^{(n)*}\mathbf{f}_{\ell'}^{(n)}\right|^2 + \frac{\sigma^2}{P_{\mathrm{d}}}}, \tag{1.16}$$

By [64, 70], similar to the uplink case, a lower bound of the downlink achievable ergodic rate is

$$\xi_{\mathrm{d}} = \mathbb{E}[\log_2(1+\mathrm{SINR}_{\mathrm{d}})]. \tag{1.17}$$

Last, I compute the average cell throughput ξ_{cell} as the sum rate of all users in a cell as follows. Let T_{u} and T_{d} be the lengths of uplink and downlink data transmission in a frame, in terms of the number of symbols. To account for the training overhead, recall that the length of the uplink training equals the length of the pilots τ. Define the length of the frame T_{f} as

$$T_{\mathrm{f}} = T_{\mathrm{u}} + T_{\mathrm{d}} + \tau. \tag{1.18}$$

Note that T_{f} should be smaller than the coherent time T_{c}. Then the average cell throughput ξ_{cell} can be computed as

$$\xi_{\mathrm{cell}} = \frac{K}{T_{\mathrm{f}}} \left(T_{\mathrm{u}} \xi_{\mathrm{u}} + T_{\mathrm{d}} \xi_{\mathrm{d}} \right). \tag{1.19}$$

1.2.5 Asymptotic Performance Analysis

Next, based on the system model, I present analytical results on the asymptotic performance of massive MIMO, when the number of antennas grows to infinity [90]. Key concepts in the performance analysis of massive MIMO, e.g. the asymptotic orthogonality of channel vectors, and the limiting factor as pilot contamination will be illustrated in this section.

To begin with, I present results on the asymptotic orthogonality of independent Gaussian vectors in the following lemma.

Lemma 1.2.1 (From [90])**.** *Let* $\mathbf{x}$, $\mathbf{x} \in \mathcal{C}^{M \times 1}$ *be independent random vectors following the distribution* $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{\mathrm{M}})$*. Then the two vectors are asymptotic orthogonal to each other in the limit of vector dimension, as*

$$\lim_{M \to \infty} \frac{\mathbf{x}^* \mathbf{y}}{M} \overset{a.s.}{=} 0. \tag{1.20}$$

In addition, the square norm of the vector $\mathbf{x}$ *converges as*

$$\lim_{M \to \infty} \frac{\mathbf{x}^* \mathbf{x}}{M} \overset{a.s.}{=} 1. \tag{1.21}$$

Proof. The proof directly follows from the strong law of large numbers. □

Note that results in Lemma 1.2.1 show that assuming independent Rayleigh fading, the channel vector for the desired link becomes orthogonal to both the noise vector and channel vector for interfering links, when the number of antennas goes to infinity. Based on Lemma 1.2.1, I can derive the asymptotic limit for the uplink SINR, when $M \to \infty$.

Theorem 1.2.1 (From [90]). *When the number of base station antennas goes to infinity, the uplink SINR in (1.13) converges to its asymptotic equivalence as*

$$\lim_{M\to\infty} \mathrm{SINR}_{\mathrm{u}} = \frac{\beta_{\ell\ell}^{(k)2}}{\sum_{\ell'\neq\ell} \beta_{\ell\ell'}^{(k)2}}. \tag{1.22}$$

Theorem 1.2.1 shows that even with infinite antennas, the uplink SINR will not goes to infinity, as certain interference terms will not vanish, due to the impact from pilot contamination. Similarly, the asymptotic downlink SINR result is presented as follows.

Theorem 1.2.2 (From [90]). *When the number of base station antennas goes to infinity, the downlink SINR in (1.16) converges to its asymptotic equivalence as*

$$\lim_{M\to\infty} \mathrm{SINR}_{\mathrm{d}} = \frac{\beta_{\ell\ell}^{(k)2}/a_{\ell}^{(k)}}{\sum_{\ell'\neq\ell} \beta_{\ell'\ell}^{(k)2}/a_{\ell'}^{(k)}}. \tag{1.23}$$

The residue interference terms in the asymptotic equivalence in (1.23) are also resulted from pilot contamination.

By now, I have introduced basic concepts of massive MIMO networks using the IID Rayleigh fading case as an example. Through the derivation of

asymptotic analysis, I have also illustrated the impact of pilot contamination and asymptotic orthogonality on the SINR and rate performance. In the next section, I are going to provide research surveys on several key aspects of massive MIMO.

1.3 Key Aspects of Massive MIMO

Now I present literature on several key directions of research on massive MIMO. I first summarize results on more complicated designs of precoder and combiner in Section 1.3.1. Then I focus on the issue of pilot contamination, and introduce methods de-contamination methods in Section 1.3.2. In Section 1.3.3, I show that massive MIMO is a promising technique for green communication by improving energy efficiency. In the end, I will briefly introduce field measurements and channel models for massive MIMO in Sections 1.3.4 and 1.3.5.

1.3.1 Precoding and Combining Techniques

In this section, I focus on linear precoding and combining techniques for massive MIMO, as linear techniques has less complexity to implement than the non-linear techniques, such as dirty paper precoding, especially when the number of antennas is large. Simulations also shows that with a large number of antennas, the achievable rate of linear precoding approaches the channel capacity bound [124].

Single-cell precoding I start with the single-cell precoding design, when there is no cooperation between base stations. Namely, base station X_ℓ designs the precoders and combiners to its users, based on the channel estimate matrix $\bar{\mathbf{H}}_{\ell\ell}$, whose k-th column $\bar{\mathbf{h}}_{\ell\ell}^{(k)}$ is the estimate of channel $\mathbf{h}_{\ell\ell}^{(k)}$.

For the uplink combining, let $\mathbf{G}_\ell$ be the combing matrix for base station X_ℓ, where its k-th column $\mathbf{g}_\ell^{(k)}$ is the combining vector for user $Y_\ell^{(k)}$. A regularized zero-forcing (RZF) combiner can be computed as

$$\mathbf{G}_\ell = \left(\bar{\mathbf{H}}_{\ell\ell}\bar{\mathbf{H}}_{\ell\ell}^* + \rho \mathrm{I}_N\right)^{-1} \bar{\mathbf{H}}_{\ell\ell}, \tag{1.24}$$

where the regularization constant ρ can be optimized based on performance requirements. When $\delta \to \infty$, the RZF combiner becomes match-filtering combiner as I used in Section 1.2.4 as

$$\mathbf{G}_\ell = \bar{\mathbf{H}}_{\ell\ell}; \tag{1.25}$$

when $\delta \to 0$, and $\bar{\mathbf{H}}_{\ell\ell}$ has a full column rank, the combiner becomes zero-forcing combiner as

$$\mathbf{G}_\ell = \left(\bar{\mathbf{H}}_{\ell\ell}\bar{\mathbf{H}}_{\ell\ell}^*\right)^{-1} \bar{\mathbf{H}}_{\ell\ell}, \tag{1.26}$$

which is the pseudo-inverse of $\bar{\mathbf{H}}_{\ell\ell}$.

For the downlink precoder, let $\mathbf{F}_\ell$ be the precoder matrix for base station X_ℓ, where its k-th column $\mathbf{f}_\ell^{(k)}$ is the percoder for user $Y_\ell^{(k)}$. A RZF beamformer can be computed as

$$\mathbf{F}_\ell = \left(\bar{\mathbf{H}}_{\ell\ell}\bar{\mathbf{H}}_{\ell\ell}^* + \rho \mathrm{I}_N\right)^{-1} \bar{\mathbf{H}}_{\ell\ell}\mathbf{D}_\ell, \tag{1.27}$$

where $\mathbf{D}_\ell \in \mathcal{R}^{K\times K}$ is a diagonal matrix, and its diagonal element $\mathbf{D}_\ell[k,k]$ is the power normalization term for user $Y_\ell^{(k)}$ to account for the constraint on power constraint. Similar to the uplink case, when $\delta \to \infty$, the RZF precoder $\mathbf{F}_\ell = \bar{\mathbf{H}}_{\ell\ell}\mathbf{D}_\ell$ becomes conjugate precoder, also known as maximum ratio transmission; when $\delta \to 0$, and $\bar{\mathbf{H}}_{\ell\ell}$ has full column rank, the RZF precoder $\mathbf{F}_\ell = \left(\bar{\mathbf{H}}_{\ell\ell}\bar{\mathbf{H}}_{\ell\ell}^*\right)^{-1}\bar{\mathbf{H}}_{\ell\ell}\mathbf{D}_\ell$ is called zero-forcing precoder.

Common methods to compute $\mathbf{D}$ are vector normalization and matrix normalization. For vector normalization, $\mathbf{D}_\ell$ is determined such that each column of $\mathbf{F}_\ell$ has a unit norm as $\|\mathbf{f}_\ell^{(k)}\|^2 = 1$; for matrix normalization, $\mathbf{D}_\ell = d_\ell \mathbf{I}_K$, where the constant d_ℓ ensures that $\|\mathbf{F}_\ell\|_{\mathrm{F}}^2 = K$. In [83], the analysis showed that vector normalization provides a better achievable rate in the low signal-to-noise ratio (SNR) regime, while matrix normalization provides a marginal gain in the high SNR regime.

In [147], the performance of RZF precoder was analyzed in a single-cell massive MIMO network with IID Rayleigh fading, where the impact of pilot contamination was not explicitly taken account. Low bounds for the channel capacity were derived for both conjugate and ZF precoders. In terms of the achievable rate, the analysis showed that ZF precoder outperforms in the high SNR regime, while the reverse is true in the low SNR regime; an (approximate) boundary SNR to switch from conjugate to ZF precoder was provided in [83].

The uplink and downlink rate performance was examined in a multi-cell setup in [64], which incorporated the impacts of pilot contamination and spatial correlations at base station antennas. Based on random matrix theory,

the uplink and downlink achievable rates for RZF combiner and precoder were expressed as functions of the number of antennas and scheduled user. The asymptotic analysis showed that when the same covariance matrix is assumed for both signal and interference links, the conjugate and RZF beamforming have the same asymptotic SINR, limited by pilot contamination. The analyses for the case of finite base station antennas, however, indicates that ZF beamforming converges fast to the asymptotic limit than conjugate beamforming.

Multi-cell precoding Multi-cell coordinated transmission and reception have been considered in massive MIMO networks. While mainly aimed to managing interference in conventional cellular networks[125], coordinated transmission in massive MIMO networks is applied to overcome the form factor constraint by reducing the number of antennas on a single base station. In [65], a network MIMO architecture was proposed to achieve the throughput of massive MIMO with a order of magnitude fewer antennas at each base station. In [137, 88], the concept of distributed massive MIMO was proposed, where the antennas are distributed over multiple remote radio heads (RRHs) in a cell, and jointly perform channel estimation and data transmission. Distributed massive MIMO can achieve even better throughput than conventional systems, e.g. simulations in [137] showed that with the same total number of antennas, distributed massive MIMO achieves a higher average uplink rate than the case of collocated antenna array. Besides overcoming the form factor constraint, multi-cell coordinated precoding have also been considered to

mitigate pilot contamination [21, 78] (see Section 1.3.2), and reduce the sum transmission power [74].

To enable multi-cell coordinated transmission in massive MIMO systems, the CSI needs to be jointly acquired and shared between base stations, which requires a larger amount of overhead compared with conventional systems [44]. One promising approach to reduce the overhead due to CSI sharing is to design coordinated beamforming algorithms based on long-term channel statistics instead of instantaneous CSI [74, 21, 78]. Besides, decentralized beamforming algorithms have been proposed to reduce the overhead of distributed massive MIMO systems [88]. For example, in [137], an RRH selection method was proposed for distributed MRC receivers by grouping users locally, and shown to approach the performance of minimum mean square error (MMSE) receivers.

Non-ideal hardware constraints In practice, large-scale arrays are likely to be built with low-cost components that are prone to hardware imperfection. Therefore, constraints due to non-ideal hardware should be taken account in the designs of massive MIMO networks. To analyze the performance of massive MIMO systems with non-ideal hardware, a general framework was proposed to model the additive distortions in [37], and extended to incorporate multiplicative phase-drifts and noise amplification in [38]. The analysis in [37] showed that hardware impairments degrade the CSI accuracy, and create a finite ceiling on the achievable rate; the hardware at the mobile station side is the main limiting factor of the capacity, while the impact of base station

hardware impairments vanishes asymptotically. Moreover, results in [37, 38] also proved that with excessive degrees of freedom offered by the large antenna arrays, massive MIMO systems become more tolerable with hardware impairments, which motivates the use of low-cost antenna elements at the base station arrays.

To improve the efficiency of the power amplifier, low peak-to-average-power-ratio (PAPR) precoding designs were proposed in [98, 132] for massive MIMO systems. For example, a per-antenna constant envelope precoding was developed in [98] to minimize the PAPR of transmit signal and allow for less power back-off at the power amplifiers.

Massive MIMO networks exploits the reciprocity between downlink and uplink links to estimation channels. In practice, however, the end-to-end channel reciprocity breaks down, as the transmitting and receiving branches at one transceiver may have different circuit gains. Based on the analysis in [154], such hardware mismatch has a larger impact on RZF beamforming than conjugate beamforming on the ergodic achievable rate. In addition, to restore the channel reciprocity, calibration methods have been proposed, e.g. in [154, 123, 139], to compensate for the mismatch in the circuit gains.

1.3.2 Pilot Contamination

Pilot contamination results from the reuse of pilot sequences from one cell to another. In a massive MIMO system, when a base station correlates its received pilot signal with the corresponding pilot sequence, it obtains a chan-

nel estimate that is contaminated by a linear combination of the interfering channels from the users sharing the same pilot [76]. Pilot contamination is not specific to massive MIMO systems, but its impact becomes more phenomenal in massive MIMO, as it becomes an limiting factor for the SINR performance with infinity antennas [90] (also see the discussion in Section 1.2.3). Therefore, several methods have been proposed to reduce or mitigate the impact of pilot contamination in massive MIMO networks. I introduce main methods by category as follows.

1. Reducing pilot reuse in a resource block One way to reduce the impact of pilot contamination is to decrease the number of users sharing the same pilot in a particular resource block. One direct thought is to increase the number of orthogonal pilot sequences in the network, such that users in the neighbouring cells need not share pilots. Increasing the number of orthogonal pilot sequences, however, results in longer sequences, and a larger amount of training overhead. More importantly, the maximum number of the orthogonal pilots is fundamentally limited by the coherence time in the system.

Similar to the idea of fraction frequency reuse to reduce interference in the cell-edge area, a fractional pilot reuse scheme was proposed to reduce pilot contamination in [22]. The fraction frequency reuse method was shown to have higher cell throughput than the full pilot reuse in [90], as the cell-edge users become less affected by pilot contamination.

Frequency reuse, i.e., assigning different bands for neighbouring cells, also reduces pilot contamination by mitigating interference from neighbouring

cells. Splitting spectrum, however, need not help improve the sum throughput, as fewer users are served in a unit spectrum [87].

Prior work showed that synchronous pilot transmission, as I assumed in Section 1.2, maximizes the effect of pilot contamination [90]. Therefore, a shifted-frame protocol was proposed in [19] to enable asynchronous pilot transmission in adjacent cells, where base stations are divided into different groups, and uplink channel training stages for cells in different groups are performed in disjoint and shifted time periods within a frame. In [19], the shifted-frame protocol was shown to achieve a better asymptotic SINR and rate performance, as pilot contamination only affects the users within the same group of cells.

2. Coordinating between multiple cells Multi-cell coordination is an effective way to reduce the impact of pilot contamination. In [150], a coordinated scheduling algorithm was proposed to minimize the estimation error due to pilot contamination from adjacent cells. The theoretical basis of the algorithm is that given the eigenvectors of the covariance matrices for the signal link and interference links span in orthogonal spaces, an interference-free channel estimate of the signal link can be obtained through certain MMSE estimator [150]. One example satisfying the orthogonal eigenvector space condition is that a uniform linear array receives the desired signal and interference in disjoint intervals of arrival angles [149, 150]. Therefore, the scheduling algorithm requires base stations share channel covariance information, and jointly pick up the users, to guarantee the users with the same pilots have minimal

overlapping in the space of covariance matrices [150].

Besides coordinated scheduling, a coordinated precoding method called pilot contamination precoding (PCP) was proposed in [21, 78] to mitigate pilot contamination. To limit the overhead of information exchange, only large-scale path losses and information symbols are shared between base stations. I explain the basic idea using the uplink ZF-PCP [21] as an example. I consider the massive MIMO network model of L cells in Section 1.2. When the number of antennas $M \gg 1$, the detected symbol $\hat{s}_\ell^{(k)}$ at base station X_ℓ can be approximated by its asymptotic equivalence as

$$\hat{s}_\ell^{(k)} \approx M\sqrt{P_{\mathrm{u}}} \sum_{\ell'=1}^{L} c_{\ell\ell'}^{(k)} s_{\ell'}^{(k)}, \tag{1.28}$$

where the approximation becomes exact when $M \to \infty$, and the coefficient $c_{\ell\ell'}^{(k)} = \beta_{\ell\ell}^{(k)}\beta_{\ell\ell'}^{(k)}/a_\ell^{(k)}$ only depends on large-scale path losses. By sharing the detected signals $\hat{s}_\ell^{(k)}$ and coefficients $c_{\ell\ell'}^{(k)}$ among all L base stations, I obtain L independent (approximate) linear combination of $\left\{s_\ell^{(k)}\right\}_{0\le\ell\le L-1}$ as

$$\begin{bmatrix} \hat{s}_0^{(k)} \\ \vdots \\ \hat{s}_{L-1}^{(k)} \end{bmatrix} \approx M\sqrt{P_{\mathrm{u}}} \begin{bmatrix} c_{00}^{(k)} & \cdots & c_{0(L-1)}^{(k)} \\ \vdots & \ddots & \vdots \\ c_{(L-1)0}^{(k)} & \cdots & c_{(L-1)(L-1)}^{(k)} \end{bmatrix} \begin{bmatrix} s_0^{(k)} \\ \vdots \\ s_{L-1}^{(k)} \end{bmatrix}. \tag{1.29}$$

The contamination-free symbols can be approximated by solving (1.29), which returns the true uplink symbol asymptotically. The downlink ZF-PCP design follows from a similar idea as in the uplink [21]. In [78], more complicated PCP precoder designs were proposed to optimize the minimal achievable rate with finite antennas. Simulations in [78] showed that the proposed method

provides an order of magnitude increase in rate with 100 antennas at 10% SINR outage .

3. Developing blind estimation algorithm Blind algorithms have been proposed to estimate the channels without using (user-specific) pilot sequences [100, 103] in the training. In general, based on subspace separation techniques, blind channel estimation algorithms rely on the assumptions that the signal subspace is asymptotically orthogonal to the subspace of interference, when the number of antennas grows without limit [100, 103]. In [103], the channel matrix is estimated through eigenvalue decomposition of the covariance matrix of the received signal. In practice, the covariance matrix of the received signal is approximated by the empirical matrix covariance from uplink data samples, as the exact covariance matrix is not unknown at base stations. Therefore, as simulation results showed in [103], the performance of the proposed blind algorithm largely depends on the accuracy in estimating the covariance matrix; with a sufficiently large number of uplink data symbols to estimate the covariance matrix, blind channel estimation outperforms the original pilot-aided channel estimation in symbol error probability.

In [100], the received signal is projected to the signal subspace, which is asymptotically orthogonal to the interference space, to minimize interference. To identify the signal subspace of interests, a certain power margin between the desired signal and interference at cell-edge users is required, which can be created by power control and frequency reuse.

1.3.3 Energy efficiency

Massive MIMO is a promising technique to improve energy efficiency, an key performance metric in green communication [148]. Energy efficiency is defined as the ratio of the spectrum efficiency to the total power, including the transmit power and consumption power by the circuit.

Energy efficiency for massive MIMO systems was examined in [104, 147], where only the transmit power was considered. In [104], the (asymptotic) power scaling laws for uplink massive MIMO systems were derived as follows. In a single-cell network, assuming IID Rayleigh fading, perfect CSI, and infinite antennas at base station, the asymptotic achievable spectrum efficiency ξ_{u} for a user is

$$\lim_{M\to\infty} \xi_{\mathrm{u}} = \log_2\left(1+\beta\frac{P_{\mathrm{u}}M}{\sigma^2}\right), \tag{1.30}$$

where β is the large-scale path loss, P_{u} is the uplink transmit power, M is the number of antennas at base station, and σ^2 is the noise power. By 1.30, the transmit power P_{u} scales with $\frac{1}{M}$ to maintain the asymptotic uplink spectrum efficiency ξ_{u} unchanged. With imperfect CSI due to thermal noise, the asymptotic achievable spectrum efficiency ξ_{u} in a single-cell network is

$$\lim_{M\to\infty} \xi_{\mathrm{u}} \approx \log_2\left(1+\tau\beta^2\frac{P_{\mathrm{u}}^2M}{\sigma^2}\right), \tag{1.31}$$

where τ is the length of pilot sequences used in the channel estimation. The error in the CSI results in a different power scaling law from the perfect CSI case: the transmit power P_{u} scales with $\frac{1}{\sqrt{M}}$ to maintain the asymptotic uplink

rate ξ_{u}. The power scaling law for the imperfect CSI case also applies to multi-cell massive MIMO systems, where the estimation error also comes from pilot contamination.

Based on the power scaling law results, the trade-off between spectrum efficiency and energy efficiency was derived in [104]. For the perfect CSI case, it is straightforward to show that the energy efficiency increases, when the spectrum efficiency decreases, as the energy efficiency is a decreasing function of P_{u}. For the imperfect CSI case, however, a non-monotonic trend was found: for small P_{u}, i.e., the low SNR regime, the energy efficiency E_{u} is

$$E_{\mathrm{u}} = \frac{\xi_{\mathrm{u}}}{P_{\mathrm{u}}} \tag{1.32}$$

$$\approx \frac{\log_2\left(1+\tau\beta^2\frac{P_{\mathrm{u}}^2 M}{\sigma^2}\right)}{P_{\mathrm{u}}} \tag{1.33}$$

$$\approx \tau\beta^2\frac{P_{\mathrm{u}}M}{\sigma^2}, \tag{1.34}$$

which increases with P_{u} and thus ξ_{u}; for large P_{u}, i.e., the high SNR regime, the energy efficiency E_{u} is

$$E_{\mathrm{u}} \approx \frac{\log_2\left(\tau\beta^2\frac{P_{\mathrm{u}}^2 M}{\sigma^2}\right)}{P_{\mathrm{u}}} \tag{1.35}$$

$$\sim \frac{\log_2 P_{\mathrm{u}}}{P_{\mathrm{u}}}, \tag{1.36}$$

which decreases with P_{u} and ξ_{u}. A similar non-monotonic trade-off trend was found in the multi-cell systems, where pilot contamination was also shown to significantly degrade performance.

In [147], the downlink energy efficiency was analyzed in a single-cell setup: conjugate beamforming outperforms ZF beamforming in the operation point of high spectrum efficiency and low energy efficiency, while the converse is true for high energy efficiency and low spectral efficiency.

In terms of circuit power consumption, full digital beamforming, which requires one RF chain per antenna, can be power consuming when the number of antennas is large [110, 81, 46]. Therefore, to reduce the circuit power consumption, more energy-efficient hardware design has been proposed in [110, 23, 81, 107, 99, 46, 142]. One approach is to use fewer RF chains than the antennas, and apply hybrid beamforming at massive MIMO base stations [23, 81, 107, 110]. Another approach is to apply low resolution analog-to-digital convertor (ADC), e.g. one-bit ADC [99, 46, 142], at receivers in the uplink to save the power due to analog-to-digital sampling.

1.3.4 Field Measurement

The performance of massive MIMO systems has been examined by prototypes and measurements in real environments [56, 57, 63, 109, 127]. In [127], a massive MIMO base station prototype called Argos was equipped with 64 antennas and capable of serving 15 users simultaneously. Experiments showed that Argos prototype can achieve up to 6.7 fold capacity gain using $\frac{1}{64}$-th of transmission power, compared with the system with a single antenna at base stations. In [57], 2.6 GHz measurement campaigns were performed using a uniform circular array and a uniform linear array, each with 128 antenna

ports, at base station side on the campus of Lund University, Sweden. The measurement showed better orthogonality between channels can be achieved using larger antenna arrays, when the users are NLOS or LOS with sufficient spatial separations. LOS channels with closed located users, however, were shown to have the worst orthogonality. This indicates that minimum separation between users in angle space should be guaranteed when scheduling LOS users. In [109], spatial correlation between base station antennas was measured using a 128 virtual uniform linear array with half-wavelength spacing at 2.6 GHz. While high correlations were observed between neighbouring antennas in both LOS and NLOS links, NLOS links generally have less variance in the correlation level.

Despite the difference between the measured channels and IID Rayleigh channels, prior work confirmed that measured channels achieve a large faction of the theoretical performance gains [56, 57, 109, 63]. For example, in [57], with 128 antennas, measured channels achieve roughly 90% of the channel capacity in IID fading channels in the NLOS links, and about half of the channel capacity in the worst case when users are LOS and closely located.

1.3.5 Channel Model

Based on the measurement results, channel modeling compromises trade-offs between the tractability and accuracy. In this section, I will focus on the analytical channel models that emphasize more on the tractability in analysis, while potential channel models for industrial massive MIMO simulations can

be found, e.g. in [49] for the WINNER+ model, [84] for the COST 2000 model, [2] for the 3GPP 3D MIMO channel model, and [156] for a survey.

One property that most analytical massive MIMO models assume is asymptotic orthogonality, which I define mathematically as follows.

Definition 1.3.1. Channel vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ are asymptotically orthogonal to each other, in terms of their dimension M, if the following limit holds in certain senses, e.g. in probability or almost surely, as

$$\lim_{M\to\infty} \frac{\mathbf{h}_1^* \mathbf{h}_2}{|\mathbf{h}_1||\mathbf{h}_2|} = 0. \tag{1.37}$$

IID fading channel model One simple yet commonly used model in the analysis is the IID Rayleigh fading model [90], where the channel vector $\mathbf{h}_\ell$ follows the distribution $\mathcal{CN}(\mathbf{0}, \beta_\ell \mathbf{I}_M)$, and the constant β_ℓ is the large-scale path loss in the ℓ-th link. As shown in Lemma 1.2.1, the IID Rayleigh fading model satisfies the property of asymptotic orthogonality. The IID fading channel models an ideal propagation environment, where the correlations in small-scaling fading is ignored.

Correlated fading channel model To model the correlated fading case, due to spatial correlations and mutual coupling, a Kronecker channel model has been used in the analysis [64], where the channel vector $\mathbf{h}_\ell$ is modeled as

$$\mathbf{h}_\ell = (\beta_\ell)^{1/2} \, \mathbf{\Phi}_\ell^{1/2} \mathbf{w}_\ell, \tag{1.38}$$

where β_ℓ is for large-scale path loss, $\mathbf{\Phi}_\ell \in \mathcal{C}^{M\times M}$ is the covariance matrix of small-scale fading, and $\mathbf{w}_\ell$ is a random vector following the distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. In this section, for ease of illustration, I denote the singular value decomposition of the $\mathbf{\Phi}_\ell$ as

$$\mathbf{\Phi}_\ell = \mathbf{U}_\ell^* \Lambda_\ell \mathbf{U}_\ell, \tag{1.39}$$

where the matrix $\Lambda_\ell \in \mathcal{R}^{M\times M}$ is a real diagonal matrix containing the singular values, and the m-th singular value is denoted as $\lambda_\ell[m] \geq 0$. In addition, I assume the matrix $\mathbf{\Phi}_\ell$ is normalized such that its trace

$$\mathrm{Tr}\left(\mathbf{\Phi}_\ell\right) = \sum_{m=1}^{M} \lambda_\ell[m] = M. \tag{1.40}$$

One example for the covariance matrix $\mathbf{\Phi}_\ell$ is the exponential correlation model [85]. In the exponential correlation model, for $m, \ell \in [1, M]$, the (m, k)-th of the matrix $\mathbf{\Phi}_\ell$ is defined as

$$\mathbf{\Phi}_\ell[m, k] = \rho_\ell^{|m-k|}, \tag{1.41}$$

where $\rho_\ell \in \mathcal{C}$ represents the correlation coefficient of fading between neighbouring antennas, and $|\rho_\ell| \in [0, 1)$.

In [150, 4], another model based on angular spread was used to characterize the the covariance matrix as

$$\mathbf{\Phi}_\ell = \mathbb{E}_{\theta_\ell}\left[\boldsymbol{\alpha}(\theta_\ell)\boldsymbol{\alpha}^*(\theta_\ell)\right], \tag{1.42}$$

where $\boldsymbol{\alpha}(\cdot) \in \mathcal{C}^{M\times 1}$ is the steering vector of the base station antenna array, θ_ℓ represents the angle of arrival for the ℓ-th link, and the expectation is taken

over the angular power spectrum to account for the angle spread. Specifically, in [150, 4] considers the case of uniform linear arrays with a single path: the steering vector $\boldsymbol{\alpha}(\cdot)$ is assumed to be

$$\boldsymbol{\alpha}(\theta) = \left[1, \mathrm{e}^{-j2\pi \frac{d\sin(\theta)}{d_\mathrm{c}}}, \mathrm{e}^{-j4\pi \frac{d\sin(\theta)}{d_\mathrm{c}}}, ...\mathrm{e}^{-j2(M-1)\pi \frac{d\sin(\theta)}{d_\mathrm{c}}}\right]', \tag{1.43}$$

where d_c is the wavelength at the carrier frequency, and d is the antenna spacing; the angular power spectrum is assumed to have a uniform distribution over a finite support $[\theta_\mathrm{c} - \delta, \theta_\mathrm{c} + \delta] \subset [0, \pi]$, where δ is the maximum angle spread. In addition, with uniform planar arrays, the 3D covariance matrix was shown to be well approximated by a Kronecker production of azimuth and elevation correlations in [151].

Depending on the structure of the covariance, correlated fading channels in 1.39 need not satisfy the asymptotic orthogonality property. One sufficient condition to guarantee the asymptotic orthogonality is given in the following theorem.

Theorem 1.3.1. *For $\ell = 1, 2$, if channel vector ℓ is modelled in 1.39, and the mean square of the eigenvalues of $\boldsymbol{\Phi}_\ell$ is uniformly bounded:*

$$\limsup_M \sum\nolimits_{m=1}^{M} \lambda_\ell^2[m]/M < \infty, \tag{1.44}$$

then the asymptotic orthogonality property in (1.37) is satisfied; as

$$\lim_{M\to\infty} \mathbf{h}_1^* \mathbf{h}_2 / M \overset{p.}{=} 0, \tag{1.45}$$

and

$$\lim_{M\to\infty} \mathbf{h}_\ell^* \mathbf{h}_\ell / M \overset{p.}{=} \beta_\ell. \tag{1.46}$$

The constraint in (1.44) is satisfied by the exponential correlation model in (1.41)[34], and the angle spread model in (1.42) with certain conditions on the angle spread range δ [4].

LOS channel model To model the extreme case with strong correlations, e.g. in a LOS link with no reflected paths, a LOS channel model has been proposed in [105, 91], where the channel vector h_ℓ is modelled by a determinist steering vector as

$$\mathbf{h}_\ell = \beta_\ell \boldsymbol{\alpha}_\ell, \tag{1.47}$$

where $\boldsymbol{\alpha}_\ell$ is often assumed to be the steering vector of a uniform linear array as defined in (1.43) [105, 91]. For LOS channel models with uniform linear arrays, one sufficient condition to achieve the asymptotic orthogonality is provided in the following theorem.

Theorem 1.3.2. *[From [105]] Let $\boldsymbol{\alpha}(\cdot)$ be the steering vector of a uniform linear array, and for $\ell = 1, 2$, θ_ℓ is the angle of arrival in the ℓ-th link. If $\sin(\theta_\ell)$ is IID uniformly distributed in $[-1, 1]$, then the asymptotic orthogonality is achieved as*

$$\lim_{M\to\infty} \frac{\boldsymbol{\alpha}(\theta_1)^* \boldsymbol{\alpha}(\theta_2)}{M} \overset{p.}{=} 0. \tag{1.48}$$

Note that the asymptotic orthogonality is not always achieved in LOS channels. In [105], one counter example was provided: when $|\sin(\theta_1) - \sin(\theta_2)| = 1/M$,

$$\lim_{M\to\infty} \frac{\boldsymbol{\alpha}(\theta_1)^* \boldsymbol{\alpha}(\theta_2)}{M} = \frac{2j}{\pi}. \tag{1.49}$$

Intuitively, the asymptotic orthogonality can be achieved only when two LOS channels are not aligned in the angular space.

Non-stationary channel models In massive MIMO systems, the dimension of the antenna array can be non-negligible compared with the link length, and such near-field effect are captured by the non-stationary channel models [143, 144, 79]. In [143, 144], the spherical wavefront at the large-scale array was modelled by an ellipse model, and the dynamics of the clusters is modelled by a birth and death process. In [79], the concept of visible region was applied to model the phenomenon that certain clusters are observed only by a fraction of the antennas in a large-scale array. Besides, the impact of channal aging, i.e., the non-stationarity in the time domain, was examined in [136], where the dynamics are modelled by an autoregressive series, and algorithms for channel prediction were also proposed to overcome channel aging.

1.4 MmWave channel statistics

The mmWave frequency ranging from 30 GHz to 300 GHz, has recently drawn great interest as a new carrier frequency for access channels in 5G cellular networks [112, 120, 17, 42]. Compared with the bandwidth-limited conventional spectrum below 6 GHz, mmWave can potentially provide several gigahertz bandwidths to solve the spectrum gridlock in cellular networks [112, 120]. Extensive measurements were conducted to have a better understanding of mmWave, e.g. see [113, 119, 120, 121] and the references therein. Though mmWave channels are dependent on site-specific environment features [119, 120], important statistics of mmWave access channels, such as the path loss laws, penetration loss, and multi-path statistics, have been derived from the measurements. In this section, we summarize the measurement results and highlight observations that we believe are important for developing analytical models to predict performance in mmWave cellular systems.

1.4.1 Penetration Loss

The penetration losses for mmWave signals through common materials were tested at 28 GHz [113], 40 GHz [113, 8], and 60 GHz [15, 75]. The measurements indicate that the penetration loss depends much on the materials: a substantial penetration loss is observed through common materials for outer walls of buildings, e.g. ¿170 dB loss through 10 cm thick brick or concrete walls [8], which indicates separate systems for indoor and outdoor coverage in mmWave bands; for the typical materials for inner walls, the penetration loss

at mmWave frequencies is similar to that in sub-6 GHz, e.g. the penetration loss through a 2.5 cm dry wall goes 5.4 dB at 2.5 GHz to 6.0 dB at 60 GHz [112], which motivates the deployment of indoor mmWave WLAN systems.

Besides buildings, mmWave signals are also attenuated by human body and trees. At mmWave frequencies, the penetration loss through human body is as high as 20-40 dB [86, 113], which increases the outage probability in mmWave cellular systems [32]. Measurements show that the foliage blocking the direct link can adds an attenuation loss of several dB [117, 102], which can be an important issue in the link budget.

1.4.2 Large-scale Path Loss

The path loss (in dB scale) in a link of length d can be modeled by the log-distance model with a close-in free space path loss as:

$$\mathrm{PL}(d) = PL(d_0) + 10\alpha \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma, \tag{1.50}$$

where $PL(d_0)$ is the close-in free space path loss, α is the path loss exponent, and X_σ is a random variable for log-normal shadowing. The close-in path loss $PL(d_0)$ can be computed by the Friis's equation as

$$\mathrm{PL}(d_0) = 10 \log_{10}\left(\frac{4\pi d_0 f_c}{v_c}\right), \tag{1.51}$$

where f_c is the carrier frequency, v_c is the speed of light, and the close-in distance is taken as $d_0 = 1$ meter in [121]. Note that the close-in path loss $PL(d_0)$ increases with the carrier frequency. When isotropic dipole antennas

are applied at both the transmitter and receiver, a mmWave signal at 30 GHz suffers a 20 dB free space path loss than that in the 3 GHz. Such difference in the path loss between mmWave and lower frequency is made up from the directivity gain of large antenna arrays. [112, 120].

Due to the existence of blockages in the direct path, measurements reveals different path loss laws for the LOS and NLOS links. In a LOS link, the path loss exponent α_{L} is found to be $\alpha_{\mathrm{L}} = 2$ in extensive measurements using directional antennas[121, 119, 113]. The LOS mmWave signals transmit as in the free space, as the direct path is less likely to be affected by the nearby obstacles, due to the smaller size of the Fresnel zone at mmWave. In the NLOS links, the path loss exponent depends on the environments, but is consistently larger than the LOS. In [121], based on measurements in the New York and Austin city, the NLOS path loss exponent was found to be $\alpha_{\mathrm{N}} = 4.5$ in Manhattan at 28 GHz, $\alpha_{\mathrm{N}} = 3.3$ in Manhattan at 38 GHz, and $\alpha_{\mathrm{N}} = 4.7$ in Manhattan at 73 GHz. Note that the exponent α_{N} depends on the scattering environment and the antenna beam width [118, 119, 120]. The reason is that the strength of the reflected signals depends on how much energy impinges on the scatterer and the type of scatterer. For example, when the antenna beam width varies from 7.8° to 49.4°, the NLOS path loss exponent decreases from 3.88 to 3.18 at 38 GHz [119], as more reflected signal power is expected to be received with an antenna with wider beam width. In addition, the shadowing variance also differs in the LOS and NLOS links, e.g. the variance is 1.1 dB for LOS, and 10.0 dB for NLOS in the 28 GHz Manhattan measurement [121].

1.4.3 Multi-path Effect and Small-scale Fading

Measurements show that a mmWave channel has fewer multi-path components than a sub-6 GHz channel, which is often called as *the channel sparsity*. For instance, the average number of multi-path components was found to be approximately 5 at 28 GHz and 3 at 73 GHz in the Manhattan area [121], while 15 clusters of components are assumed for sub-6 GHz channel in urban areas in the 3GPP and WINNER model [49]. Thanks to the sparsity in the multi-paths, compressed sensing based algorithms have been applied to estimate the channels and design beamforming, e.g. [36, 12]. Besides, due to the channel sparsity, small-scale fading has a minor impact on mmWave signals, compared with the Rayleigh fading in the sub-6 GHz channel. The results also indicate high spatial correlation in the mmWave channel and little change in received power due to small-scale fading when highly directional antennas at both the receiver and transmitter are pointing towards the boresight direction [120].

The multi-path components come with different time delays and arrival angles. We introduce the temporal and spatial statistics of the multi-path components as follows.

Delay spread Delay spread measures the difference in the arrival time of multi-paths. One commonly used metric for delay spread is the root-mean-square (RMS) delay spread, which is the empirical standard deviation of the delays for the measured multi-path components. Due to the high path loss at

higher frequencies, mmWave signals generally have smaller delay spread than the sub-6 GHz signals. Of course, the symbol period is much smaller due to the larger bandwidth in mmWave systems, so the smaller delay spread may still create intersymbol interference that requires equalization at the receiver.

The delay spread is revealed to be much dependent on the scattering environment. For instance, the delay spread was found to be larger in the New York city than in Austin, due to the more reflective environment [121]. The use of directional beamforming at the transmitter and receiver further reduces the effect of delay spread by focusing on fewer paths which arrives within the narrow beamwidth. In a LOS link, the delay spread is negligible with beam alignment on the boresight direction, and around 50 ns in the case of misalignment [119, 121]. The delay spread can be larger in the NLOS case, e.g. an average of around 10 ns at 73 GHz, even when the beams are aligned to the strongest signal direction. In additional, the results in [119] also illustrate that the delay spread may decrease with the distance between the transmitter and receiver, as fewer dominant paths arrive due to the high path loss as the transmitter-receiver separation increases.

Angle spread Measurements show that the energy of mmWave signals generally concentrates in a few principle angles of arrival (AoAs) and angles of departure (AoDs) directions [121]. Due to the scattering at the surface of the reflectors, the energy also spreads around those principle angles with certain spread, which is measured by the RMS angle spread [75]. Therefore, in the

angular space, the energy arrives and departs in a few *spatial lobes* as plotted in [120, 121]. For instance, the average angle spread (width) for a spatial lobe was 6.8° at 28 GHz and 3.7° at 73 GHz in the New York city measurements [121].

The angle spread increases and becomes more variable as the height of the receiver decreases [119, 120]. The smaller angle spread at the BSs indicates that future mmWave base stations may benefit from an adaptive array composed of a large number of antennas with somewhat narrow beamwidths [119]. Because of the larger angle spread at the receiver, it was suggested in [113] that a relatively wider beam antenna be deployed at the mobile station to capture more power.

1.4.4 Implications on MmWave System Design

Based on the channel statistics from measurements, we conclude with the following implications on the design and analysis of mmWave cellular systems.

- Indoor-outdoor penetration: With more than 170 dB penetration losses, the outer walls of certain buildings seems impenetrable for mmWave signals. Therefore, fast handover techniques between indoor and outdoor base stations are essential to guarantee the seamless coverage.

- Body and foliage blocking: One potential solution to overcome them is by exploiting macro-diversity, e.g. connecting to multiple base stations

to allow fast switch to unblocked links.

- Path loss: Different path loss laws should be applied to LOS and NLOS paths. The parameters of the models should be selected based on the specific environment, the carrier frequency, and the antenna beamwidth deployed in the system.

- Channel sparsity: MmWave channels are sparse in terms of multi-paths. Such sparsity should be leveraged to design channel estimation and beamforming algorithms. Analytical tools, e.g. the virtual channel representation model as introduced below, have been also developed to analyze the performance in mmWave systems.

- Small-scale fading: The Rayleigh fading model does not apply to mmWave systems using directional beamforming. In mmWave analysis, it is reasonable to neglect the small-scale fading when the directional antennas at both base stations and mobile station are appropriately steered. Another popular approach is to assume a general Nakagami small-scale fading, with parameters determined from the environment.

- Angle spread: As angle spread tends to be larger at the mobile stations than the base stations, a wider beam antenna array should be deployed at the mobile stations.

1.5 MmWave Signal Processing

Deploying large antenna arrays at the base stations and mobile stations is a key feature of mmWave cellular systems. Due to the use of large numbers of antennas and the constraints on hardware and power consumption, mmWave systems will deploy different transceiver architectures from the conventional MIMO transceivers at sub-6 GHz frequencies. Therefore, new MIMO signal processing techniques are required to enable mmWave commutations. In this section, we first explain how the hardware and power constraints renders different MIMO transceiver architectures in mmWave bands in Section 1.5.1. Then, we briefly introduce the signal processing techniques related to two types of potential architectures for mmWave: analog beamforming in Section 1.5.2, hybrid precoding in Section 1.5.3, and receivers with low resolution analog-to-digital converters (ADCs) in Section 1.5.4, based on the survey paper [122]. Finally, we summarize the implication of those signal process techniques to mmWave system performance in Section 1.5.5.

1.5.1 Constraints for MmWave Architectures

In conventional cellular systems below 6 GHz, the MIMO precessing is mostly considered to be performed in the digital baseband, e.g. the massive MIMO system introduced in Section 1.3. Compared with signal processing in the analog domain, digital signal processing enables a full control over both the amplitudes and phases of the entries in the precoding/combining matrices, which facilitates the implementation of sophisticated signal processing

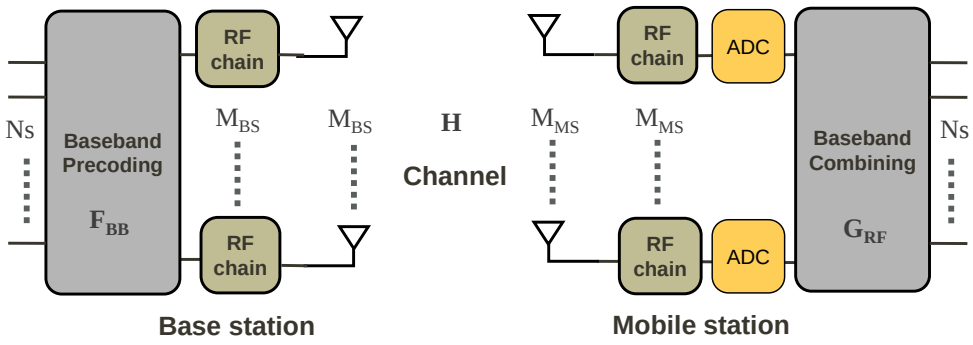


Figure 1.2: Transceiver architecture using digital precoding/combining. We illustrate the downlink transmission case as an example. Fully digital signal processing requires dedicating one RF chain per antennas.

algorithms. As shown in Fig. 1.2, however, fully digital MIMO processing requires that the transceiver dedicates an radio-frequency (RF) chain per antenna. Besides, in the Sub-6 GHz systems, the analog-to-digital quantization error tends to be a minor issue, as high-resolutions ADCs are generally assumed. With large numbers of antennas, mmWave cellular systems will not afford to deploy the transceiver architecture with full digital processing, due to the high cost and power consumption of mixed-signal components in the RF chains [52, 131, 101]. Therefore, alternative transceiver architectures have been proposed for mmWave systems [131, 52, 95, 11, 93].

One solution to address the hardware and power constraints in mmWave systems is to apply an analog/hybrid architecture that requires fewer RF chains than the antennas, and perform (part) of MIMO processing using low-cost and power-saving analog components [52, 11, 93]. Another approach is

to apply low-resolution ADCs, e.g. the ADCs with one-bit output, instead of the high-resolution ones that constitute as a main source of power consumption due to the high sampling rate required by the large mmWave bandwidth [131, 95]. We will introduce these mmWave transceiver architectures in the subsequent sections.

1.5.2 Analog Beamforming

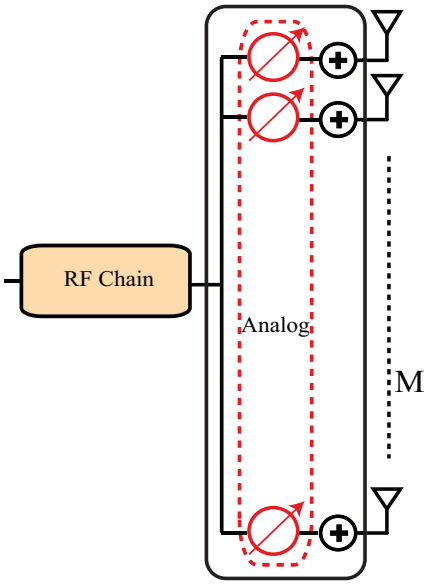


Figure 1.3: Analog transceiver architecture. Only one RF chain is required, as signal processing is performed entirely in the analog domain.

Analog beamforming is performed entirely in the analog domain, which requires only one RF chain. Due to its simplicity, it is the defacto approach in the IEEE 802.11ad standard [68]. As shown in Fig. 1.3, one common configuration to implement analog beamforming is using a network of connected phase shifters [122]. According to the channel state information, the weights

of the phase shifters can be tuned to shape and steer the beams and maximize certain performance metrics, e.g. the SNR of the received signal. With analog architecture, channel state information can be obtained via beam training with pre-designed codebooks [68, 141, 67].

Compared with digital beamforming, the performance of analog beamforming is limited, as the amplitudes of the phase shifts are constrained to be constant, and their phases can only take values from certain quantized values. More importantly, with only one RF chain, analog beamforming does not extend to transmit multiple streams or serve multiple users simultaneously. Therefore, the hybrid precoding/combining solution that use multiple RF chains (but fewer than antennas) has been proposed to overcome the limit of analog beamforming, as discussed in the next section.

1.5.3 Hybrid Analog/digital Architecture

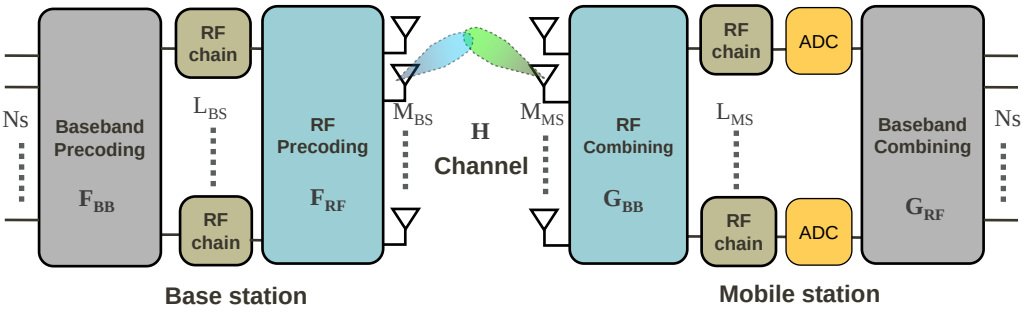


Figure 1.4: Hybrid architecture for mmWave transceivers. The number of the RF chains L should be no smaller than the number of data stream N_s, but smaller than the number of antennas M.

Hybrid analog/digital architectures divide the MIMO processing between analog and digital domain as illustrated in Fig. 1.4 [155, 52, 61, 93]. The number of RF chains L ranges between the number of data streams N_s and the number of antennas M, as a tradeoff between the hardware complexity and system performance. With $L > 1$ RF chains, the hybrid architecture allows the implementation of spatial multiplexing and multi-user MIMO.

In the hybrid architecture, the analog processing can be implemented using a network of analog components like phase shifts [52, 9], switches [93], or lens [43]. The phase shifters can be digitally controlled to provide outputs of quantized phase shifts, which can be applied to improve the precision in the analog processing. With a network of phase shifters, the RF precoder/combiner can be designed to approach the unconstrained (digital) solution [52, 9]. Using a network of switchers, antenna selection algorithms can be realized that have an even lower hardware complexity and power consumption than the phase shifters [93]. The analog processing can also be realized using lens that equivalent act as beamforming vectors with a DFT structure by performing the spatial Fourier transform [43].

There are two types of hybrid structures: (i) full connection structure [52, 9, 93] and (ii) sub-array structure [93]. In the full connection structure, a RF chain is connected to all the antennas through the network of analog components. In the second structure, the array are divided into sub-arrays, and each sub-array is fed by its own RF chain. In brief, the sub-array structure achieves a lower hardware complexity at the expense of less flexibility in the

precoder/ combiner design [61].

The design of hybrid precoder/ combiner is more challenging than in the digital domain. For one thing, the analog and digital precoding/ combining matrices need to be jointly optimized under different conditions; for another, similar to the analog beamforming case, the entries in the analog precoding/ combining matrices are subject to hardware constraints, e.g. having constant magnitudes and quantized phases [52, 9, 13, 61, 107, 153]. Thanks to the sparsity in the multi-path components, matching-pursuit based hybrid precoding designs were proposed to minimize the difference from the unconstrained (optimal) precoder [52, 9]. Other hybrid design approaches can be found in [13, 93, 107, 153]. The hybrid architectures have been shown to perform close to the constrained solutions with fully digital processing in mmWave systems with less complexity and power [52, 9, 122, 93, 61, 107, 153].

Channel estimation using hybrid architectures will be much different from the conventional approaches with fully digital processing, as training data is sent through the analog precoder/ combiner, and the entries of the channel vectors can not be accessed directly [122]. In mmWave systems, given the sparsity in the multi-paths, one promising approach is to formulate channel estimation as a sparse recovery problem that can be solved by the compressed sensing algorithms [114, 9]. For instance, in [9], an adaptive compressed sensing based approach was proposed for hybrid architectures: with multiple RF chains, the beams in the channel estimation can be shaped to approximate certain sector patterns to detect signals in the desired angular ranges, and

the angular resolution of the codebooks in each step is adaptively adjusted to reduce training overhead [9, 10].

1.5.4 Receivers With Low Resolution ADCs

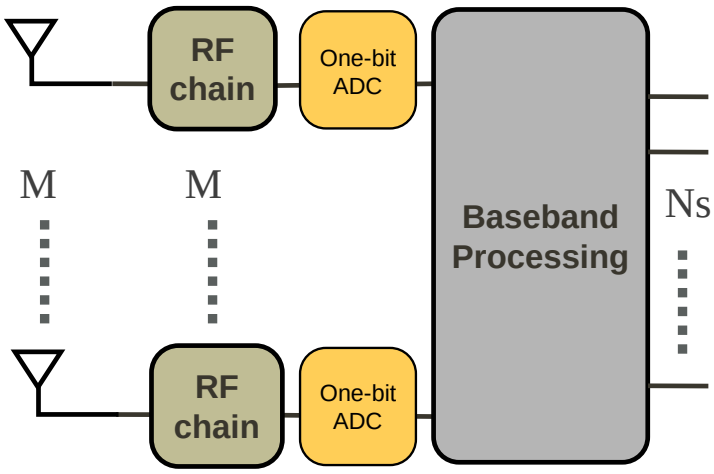


Figure 1.5: Receivers with low-resolution ADCs. Low-resolution ADCs, e.g. those with one bit output, are applied to reduce power consumption.

Due to the large system bandwidth, mmWave receivers require a high sampling rate for the ADCs. The high-rate high-resolution ADCs, however, can be power-hungry, as the power consumption of ADCs increases with the sampling rate and (exponentially) with the output bits [140]. To address the power constraints, an alternative approach to reducing the number of RF chains is to apply low-resolution ADCs at the receivers. For instance, one-bit

ADCs have drawn great interests recently, due to their low cost and power consumption [128, 96, 95].

With one-bit ADCs, channel capacity and the optimum constellation become different from those assuming infinite-bit ADCs. For instance, channel capacity will saturate to a constant at high SNR [96], while at low SNR, a 1.96 dB gap between one-bit and infinite bit ADCs was shown in [94, 95]. Moreover, channel estimation becomes even challenging with low-resolution ADCs, as the information obtained from each measurement is limited by the number of output bits. In [97], it showed that the number of measurements can be effectively reduced by applying one-bit compressed sensing algorithms, e.g. the GAMP algorithms [115], that leverage mmWave channel sparsity.

The extension from one-bit ADCs to few-bit ADCs that achieve better performance at the expense of consuming more power is still an open problem. Another interesting topic is to combine the hybrid precoding architecture with low-resolution ADCs.

1.5.5 Implications on MmWave System Design

Now based on the discussion on mmWave signal processing, we summarize some of their key implications on the system design and performance.

1. Deploying large antenna arrays at the transmitter and receiver is one key feature of mmWave cellular systems. Due to the hardware and power constraints, the MIMO transceiver architecture will be different from the conventional systems that perform fully digital processing.

2. Analog/ hybrid analog-digital precoding/ combining that uses (much) fewer RF chains than the antennas are promising for mmWave cellular systems. Due to the different transceiver architectures, the precoding/ combining and channel estimation in mmWave systems will differ from those discussed in Section 1.3 for sub-6 GHz systems.

3. Leveraging the channel sparsity in multi-path components is essential to reduce the complexity in the precoder/ combining design and overhead in channel estimation in mmWave systems.

4. Due to the high sampling rate required by the large bandwidth, mmWave receivers will likely to use ADCs with low resolutions. The impacts of quantization errors on precoding, channel estimation, and system capacity are yet to be fully understood.

1.6 Summary of Contributions

In this dissertation, I analyze the SINR and rate performance in (single-user) mmWave, sub-6 GHz massive MIMO, and mmWave massive MIMO networks. Leveraging concepts from stochastic geometry, I establish analytical frameworks that incorporate key features of these cellular networks. Analytical expressions for the SINR and rate distributions are derived in each case. Key design insights for 5G cellular networks are obtained based on the analytical results. The main contributions in each chapter are summarized as follows.

1. **Chapter 2: Analysis of MmWave Cellular Networks**

   In this chapter, I extend the sub-6 GHz stochastic geometry model to mmWave cellular networks by incorporating key mmWave features: the sensitivity to blockages and the use of large antenna arrays. Based on the model, the SINR and rate distributions are analyzed as a function of the beamforming parameters, base station density, and blockage parameters.

   (a) I develop a stochastic geometry framework that incorporates mmWave features. I use the concept of the LOS probability function, determined by the blockage parameters, to stochastically differentiate LOS and NLOS links; I apply different path loss laws to LOS/ NLOS links to incorporate the blockage effect. I approximate the beamforming pattern as a sectored antenna pattern to incorporate directional beamforming into system-level analysis.

(b) I derive analytical expressions for the SINR and rate in a downlink mmWave cellular network. I show that the SINR coverage probability in mmWave networks is much sensitive to the base station density, due to the blockage effects. In urban environments with dense buildings, a high base station density is required to avoid major coverage holes due to building blockages, and achieve an comparable SINR coverage with the conventional networks. Due to the large bandwidth at mmWave frequencies, the comparable SINR translate into a large gain in rate over the sub-6 GHz cellular networks.

(c) I demonstrate that with an increasing base station density, mmWave networks will transit from power-limited regime into interference-limited regime, where the boundary of the transition depends on the bandwidth, building distributions, and beamforming parameters (mainly beamwidth).

(d) I develop a LOS ball model to analyze the performance in dense networks. I prove that over-densification of base stations need not improve SINR, as increasing base station density also increases the likelihood to be interfered by strong LOS interferers.

2. **Chapter 3: Analysis of Uplink Massive MIMO Networks**

In this chapter, I apply stochastic geometry to analyze the uplink SINR and rate in a sub-6 GHz massive MIMO network. I incorporate pilot

contamination, a limiting factor of massive MIMO, and fractional powre control into the analysis. I consider the MRC and ZF receivers, and derive the uplink SINR and rate distributions as a function of the number of antennas and users per cell.

(a) I develop a tractable system model for uplink massive MIMO networks. To model the uplink topology, I incorporate the location correlations between the scheduled users and base stations, which make the exact analysis intractable, with a simple exclusion ball model. I use numerical simulations to show that the proposed model has a good characterization of the scheduled user process.

(b) I characterize the SINR distribution as a function of system parameters, mainly the number of antennas, the number of scheduled users per cell, base station density, and noise power. I show in simulations that with large antenna arrays at base stations, the uplink in the typical urban micro-cell case tends to be interference-limited.

(c) I study the scaling law between number of antennas and scheduled users per cell to maintain the uplink SIR distribution. I show that for MRC receivers, the SIR distribution is unchanged, when the number of antennas super-linearly scales with the number of users, unless with full channel inversion in the power control; a linear scaling law maintains the SIR distribution for ZF receivers, regardless of the power control parameters.

3. **Chapter 4: Analysis of MmWave Massive MIMO Networks**

   In this chapter, I extend the system model in Chapter 3 to mmWave frequencies, and analyze the performance of mmWave massive MIMO networks. I study the SINR in the asymptotic regime, when the number of antennas goes to infinity. I show that the asymptotic performance is a good approximation of certain practical deployments, e.g. with more than 256 antennas, when the base stations is densely deployed.

   (a) I extend the sub-6 GHz massive MIMO to model mmWave massive MIMO networks by incorporating the following features of mmWave massive MIMO: (i) the low rank and deterministic LOS channel, and (ii) directional beamforming at mobile stations.

   (b) I compute the expressions in both the downlink and uplink, when the number of antennas goes to infinity, and derive their distributions. I use numerical simulations to show that when the base station density is sufficiently dense, the SINR distribution approach the derived asymptotic limit only with 256 antennas in certain cases.

   (c) I show that in terms of cell throughput, the optimum choice of carrier frequency for massive MIMO depends on the base station density, where mmWave outperforms sub-6 GHz massive MIMO with dense base station deployments; the reverse may be true, due to blockage effect at mmWave, when the base stations are sparsely located.

1.7 Notation

I use the following notation throughout this dissertation: Bold lower-case letters $\mathbf{x}$ are used to denote vectors, and bold upper-case letters $\mathbf{X}$ are used to denote matrices. I use $\mathbf{X}[:, k]$ to denote the k-th row of matrix $\mathbf{X}$, and $\mathbf{X}^{\dagger}$ the pseudo-inverse of $\mathbf{X}$. I use $\mathbb{E}$ to denote expectation, and $\mathbb{P}$ to denote probability. $\mathcal{C}^{M \times N}$ is the $M \times N$-dimensional complex space. Other notation will be defined where needed.

Chapter 2

Analysis of MmWave Cellular Networks

2.1 Motivation and Related work

The large available bandwidth at mmWave frequencies makes them attractive for fifth generation cellular networks [116, 112, 120]. The mmWave band ranging from 30 GHz to 300 GHz has already been considered in various commercial wireless systems including IEEE 802.15.3c for personal area networking [35], IEEE 802.11ad for local area networking [68], and IEEE 802.16.1 for fixed-point access links [69]. Recent field measurements reveal the promise of mmWave signals for the access link (between the mobile station and base station) in cellular systems [119, 120].

One differentiating feature of mmWave cellular communication is the use of antenna arrays at the transmitter and receiver to provide array gain. As the wavelength decreases, antenna sizes also decrease, reducing the antenna aperture. For example, from the Friis free-space equation [59], a mmWave signal at 30 GHz will experience 20 dB larger path loss than a signal at 3 GHz. Thanks to the small wavelength, however, it is possible to pack multiple antenna elements into the limited space at mmWave transceivers [116]. With large antenna arrays, mmWave cellular systems can implement beamforming

at the transmitter and receiver to provide array gain that compensates for the frequency-dependent path loss, overcomes additional noise power, and as a bonus also reduces out-of-cell interference [112].

Another distinguishing feature of mmWave cellular communication is the propagation environment. MmWave signals are more sensitive to blockage effects than signals in lower-frequency bands, as certain materials like concrete walls found on building exteriors cause severe penetration loss [8]. This indicates that indoor users are unlikely to be covered by outdoor mmWave base stations. Channel measurements using directional antennas [113, 119, 120] have revealed another interesting behavior at mmWave: blockages cause substantial differences in the LOS paths and NLOS path loss characteristics. Such differences have also been observed in prior propagation studies at ultra high frequency bands (UHF) from 300 MHz to 3 GHz, e.g. see [1]. The differences, however, become more significant for mmWave since diffraction effects are negligible [112], and there are only a few scattering clusters [6]. Measurements in [113, 119, 120] showed that mmWave signals propagate as in free space with a path loss exponent of 2. The situation was different for NLOS paths where a log distance model was fit with a higher path loss exponent and additional shadowing [119, 120]. The NLOS path loss laws tend to be more dependent on the scattering environment. For example, an exponent as large as 5.76 was found in downtown New York City [120], while only 3.86 was found on the UT Austin campus [119]. The distinguishing features of the propagation environment need to be incorporated into the any comprehensive system analysis of

mmWave networks.

The performance of mmWave cellular networks was simulated in prior work [6, 5] using insights from propagation channel measurements [120]. In [5], using the NLOS path loss law measured in the New York City, lower bounds of the SINR distribution and the achievable rate were simulated in a 28 GHz pico-cellular system. In [6], a mmWave channel model that incorporated blockage effects and angle spread was proposed and further applied to simulate the mmWave network capacity. Both results in [6, 5] show that the achievable rate in mmWave networks can outperform conventional cellular networks in the ultra high frequency (UHF) band by an order-of-magnitude. The simulation-based approach [5, 6] does not lead to elegant system analysis as in [18], which can be broadly applied to different deployment scenarios.

Stochastic geometry is a useful tool to analyze system performance in conventional cellular networks [18]. In [18], by modeling base station locations in a conventional cellular network as a Poisson point process (PPP) on the plane, the aggregate coverage probability was derived in a simple form, e.g. a closed-form expression when the path loss exponent is 4. Moreover, the stochastic model was shown to provide a lower bound of the performance in a real cellular system [18]. There have been several extensions of the results in [18], such as analyzing a multi-tier network in [50] and predicting the site-specific performance in heterogeneous networks in [62]. It is not possible to directly apply results from conventional networks to mmWave networks due to the different propagation characteristics and the use of directional beam-

forming. There has been limited application of stochastic geometry to study mmWave cellular networks. The primary related work was in [7], where directional beamforming was incorporated for single and multiple user configurations, but a simplified path loss model was used that did not take mmWave propagation features into account.

A systematic study of mmWave network performance should incorporate the impact of blockages such as buildings in urban areas. One approach is to model the blockages explicitly in terms of their sizes, locations, and shapes using data from a geographic information system. This approach is well suited for site-specific simulations [126] using electromagnetic simulation tools like ray tracing [135]. An alternative is to employ a stochastic blockage model, e.g. [55, 30], where the blockage parameters are drawn randomly according to some distribution. The stochastic approach lends itself better to system analysis and can be applied to study system deployments under a variety of blockage parameters such as size and density.

2.2 Contributions

The main contribution of this chapter is to propose a stochastic geometry framework for analyzing the coverage and rate in mmWave cellular networks. As a byproduct, the framework also applies to analyze heterogenous networks in which the base stations are distributed as certain non-homogeneous PPPs. We incorporate directional beamforming by modeling the beamforming gains as marks of the base station PPPs. For tractability

of the analysis, the actual beamforming patterns are also approximated by a sectored model, which characterizes key features of an antenna pattern: directivity gain, half-power beamwidth, and front-back ratio. A similar model was also employed in work on ad hoc networks [66]. To incorporate blockage effects, we model the probability that a communication link is LOS as a function of the link length, and provide a stochastic characterization of the region where a user does not experience any blockage, which we define as the *LOS region*. Applying the distance-dependent LOS probability function, the base stations are equivalently divided into two independent non-homogenous point processes on the plane: the LOS and the NLOS base station processes. Different path loss laws and fading are applied separably to the LOS and NLOS case. Based on the system model, expressions for the SINR and rate coverage probability are derived in general mmWave networks. To simplify the analysis, we also propose a systematic approach to approximate a complicated LOS function as its equivalent step function. Our analysis indicates that the coverage and rate are sensitive to the density of base stations and the distribution of blockages in mmWave networks. It also shows that dense mmWave networks can generally achieve good coverage and significantly higher achievable rate than conventional cellular networks.

A simplified system model is proposed to analyze dense mmWave networks, where the infrastructure density is comparable to the blockage density. For a general LOS function, the LOS region observed by a user has an irregular and random shape. Coverage analysis requires integrating the SINR over this

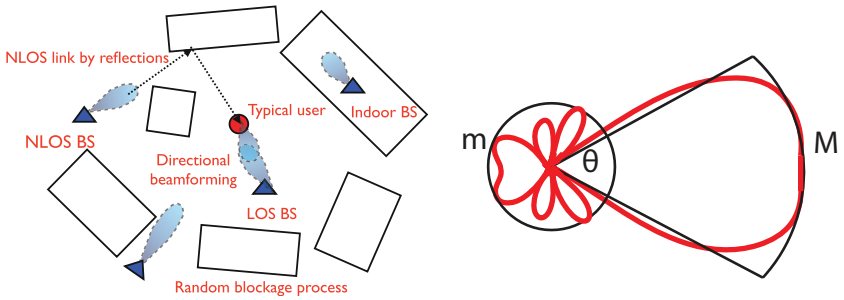
region [26]. We propose to simplify the analysis by approximating the actual LOS region as a fixed-sized ball called the *equivalent LOS ball*. The radius of the equivalent LOS ball is chosen so that the ball has the same average number of LOS base stations in the network. With the simplified network model, we find that in a dense mmWave network, the cell radius should scale with the size of LOS region to maintain the same coverage probability. We find that continuing to increase base station density (leading to what we call ultra-dense networks) does not always improve SINR, and the optimal base station density should be finite.

Compared with our prior work in [26], this chapter provides a generalized mathematical framework and includes the detailed mathematical derivations. The system model applies for a general LOS probability function and includes the impact of general small-scale fading. We also provide a new approach to compute coverage probability, which avoids inverting the Fourier transform numerically and is more efficient than prior expressions in [26]. Compared with our prior work in [31], we also remove the constraint that the LOS path loss exponent is 2, and extend the results in [31] to general path loss exponents, in addition to providing derivations for all results, and new simulation results.

This chapter is organized as follows. We introduce the system model in Section 2.3. We derive expressions for the SINR and rate coverage in a general mmWave network in Section 2.4. A systematic approach is also proposed to approximate general LOS probability functions as a step function to further

simplify analysis. In Section 2.5, we apply the simplified system model to analyze performance and examine asymptotic trends in dense mmWave networks, where outdoor users observe more than one LOS base stations with high probability. Finally, conclusions are provided in Section 2.7.

2.3 System Model



(a) System model for mmWave cellular networks

(b) Sectored model to approximate beamforming patterns.

Figure 2.1: In (a), we illustrate the proposed system model for mmWave cellular networks. Blockages are modeled as a random process of rectangles, while base stations are assumed to be distributed as a Poisson point process on the plane. An outdoor typical user is fixed at the origin. The base stations are categorized into three groups: indoor base stations, outdoor base stations that are LOS to the typical user, and outdoor base stations that are NLOS to the user. Directional beamforming is performed at both base stations and mobile stations to exploit directivity gains. In (b), we illustrate the sectored antenna model $G_{M,m,\theta}(\phi)$, which is used to approximate the beamforming patterns.

In this section, we introduce our system model for evaluating the performance of a mmWave network. We focus on downlink coverage and rate experienced by an outdoor user, as illustrated in Fig. 2.1(a). We make the following assumptions in our mathematical formulation.

Assumption 2.3.1 (Blockage process)**.** *The blockages, typically buildings in urban areas, form a process of random shapes, e.g. a Boolean scheme of rectangles [30], on the plane. We assume the distribution of the blockage process to be stationary and isotropic - in other words - invariant to the motions of translation and rotation [24, Chapter 10].*

Assumption 2.3.2 (PPP BS)**.** *The base stations form a homogeneous PPP $\tilde{\Phi}$ with density $\tilde{\lambda}$ on the plane. Note that a base station can be located either inside a blockage or outside a blockage. In this chapter, however, we will focus on the SINR and rate provided by the outdoor base stations as the blockages are assumed to be impenetrable. Let $\Phi = \{X_\ell\}$ be the point process of outdoor base stations, X_ℓ the ℓ-th outdoor base station, and $R_\ell = |OX_\ell|$ denote the distance from ℓ-th base station to the origin O. Define τ as the average fraction of the land covered by blockages, i.e., the average fraction of indoor area in the network. Further, we assume the base station process $\tilde{\Phi}$ is independent of the blockage process. Therefore, each base station has an i.i.d. probability $1-\tau$ to be located outdoor. By the thinning theorem of PPP [24], the outdoor base station process Φ is a PPP of density $\lambda = (1-\tau)\tilde{\lambda}$ on the plane. In addition, all base stations are assumed to have a constant transmit power P_{t}.*

Assumption 2.3.3 (Outdoor user)**.** *The users are distributed as a stationary point process independent of the base stations and blockages on the plane. A typical user is assumed to be located at the origin O, which is a standard approach in the analysis using stochastic geometry [18, 24]. By the stationarity and independence of the user process, the downlink SINR and rate experienced*

by the typical user have the same distributions as the aggregate ones in the network. The typical user is assumed to be outdoors. The indoor-to-outdoor penetration loss is assumed to be high enough such that an outdoor user can not receive any signal or interference from an indoor base station. Therefore, the focus in this chapter is on investigating the conditional SINR and rate distribution of the outdoor typical user served by outdoor infrastructure. Indoor users can be served by either indoor base stations or by outdoor base stations operated at UHF frequencies, which have smaller indoor-to-outdoor penetration losses in many common building materials. We defer the extension to incorporate indoor users to future work.

We say that a base station at X is *LOS* to the typical user at the origin O if and only if there is no blockage intersecting the link OX. Due to the presence of blockages, only a subset of the outdoor base stations Φ are LOS to the typical user.

Assumption 2.3.4 (LOS and NLOS BS)**.** *An outdoor base station can be either LOS or NLOS to the typical user. Let Φ_{L} be the point process of LOS base stations, and $\Phi_{\mathrm{N}} = \Phi/\Phi_{\mathrm{L}}$ be the process of NLOS base stations. Define the LOS probability function $p(R)$ as the probability that a link of length R is LOS. Noting the fact that the distribution of the blockage process is stationary and isotropic, the LOS probability function depends only on the length of the link R. Also, $p(R)$ is a non-increasing function of R; as the longer the link, the more likely it will be intersected by one or more blockages. The NLOS probability of a link is $1 - p(R)$.*

The LOS probability function in a network can be derived from field measurements [6] or stochastic blockage models [30, 55], where the blockage parameters are characterized by some random distributions. For instance, when the blockages are modeled as a rectangle boolean scheme in [30], it follows that $p(R) = \mathrm{e}^{-\beta R}$, where β is a parameter determined by the density and the average size of the blockages, and $1/\beta$ is what we called the average LOS range of the network in [30].

For the tractability of analysis, we further make the following independent assumption on the LOS probability; taking account of the correlations in blockage effects generally makes the exact analysis difficult.

Assumption 2.3.5 (Independent LOS probability)**.** *The LOS probabilities are assumed to be independent between different links, i.e., we ignore potential correlations of blockage effects between links.*

Note that the LOS probabilities for different links are not independent in reality. For instance, neighboring base stations might be blocked by a large building simultaneously. Numerical results in [30], however, indicated that ignoring such correlations cause a minor loss of accuracy in the SINR evaluation. Assumption 2.3.5 also indicates that the LOS base station process Φ_{L} and the NLOS process Φ_{N} form two independent non-homogeneous PPP with the density functions $p(R)\lambda$ and $(1-p(R))\lambda$, respectively, where R is the radius in polar coordinates.

Assumption 2.3.6 (Path loss model)**.** *Different path loss laws are applied to*

LOS and NLOS links. Given a link has length R, its path loss gain $L(R)$ is computed as

$$L(R) = \mathbb{I}(p(R))C_{\mathrm{L}}R^{-\alpha_{\mathrm{L}}} + (1 - \mathbb{I}(p(R))C_{\mathrm{N}}R^{-\alpha_{\mathrm{N}}}, \tag{2.1}$$

where $\mathbb{I}(x)$ is a Bernoulli random variable with parameter x, α_{L}, α_{N} are the LOS and NLOS path loss exponents, and C_{L}, C_{N} are the intercepts of the LOS and NLOS path loss formulas. Typical values of mmWave path loss exponents and intercept constants are available in prior work, see e.g. [119, 120]. The model could be further enhanced by including log-normal shadowing, but this is deferred in our chapter to simplify the analysis.

Assumption 2.3.7 (Directional beamforming)**.** *Antenna arrays are deployed at both base stations and mobile stations to perform directional beamforming. For tractability of the analysis, the actual array patterns are approximated by a sectored antenna model, which was used in prior ad hoc network analysis [66]. Let $G_{M,m,\theta}(\phi)$ denote the sectored antenna pattern in Fig. 1(b), where M is the main lobe directivity gain, m is the back lobe gain, θ is the beamwidth of the main lobe, and ϕ is the angle off the boresight direction. In the sectored antenna model, the array gains are assumed to be constant M for all angles in the main lobe, and another constant m in the side lobe in the sectored model. We let M_{t}, m_{t}, and θ_{t} be the main lobe gain, side lobe gain, and half power beamwidth of the base station antenna, and M_{r}, m_{r}, and θ_{r} the corresponding parameters for the mobile station. Without loss of generality, we denote the boresight direction of the antennas as $0°$. Further, let $D_\ell = G_{M_{\mathrm{t}},m_{\mathrm{t}},\theta_{\mathrm{t}}}(\phi_{\mathrm{t}}^\ell)G_{M_{\mathrm{r}},m_{\mathrm{r}},\theta_{\mathrm{r}}}(\phi_{\mathrm{r}}^\ell)$ be the total*

Table 2.1: Probability Mass Function of D_ℓ and $\bar{D}_\ell$

k	1	2	3	4
a_k	$M_\mathrm{r}M_\mathrm{t}$	$M_\mathrm{r}m_\mathrm{t}$	$m_\mathrm{r}M_\mathrm{t}$	$m_\mathrm{r}m_\mathrm{t}$
b_k	$c_\mathrm{r}c_\mathrm{t}$	$c_\mathrm{r}(1-c_\mathrm{t})$	$(1-c_\mathrm{r})c_\mathrm{t}$	$(1-c_\mathrm{r})(1-c_\mathrm{t})$
e_k	M_r	$M_\mathrm{r}/\xi_\mathrm{t}$	m_r	$m_\mathrm{r}/\xi_\mathrm{t}$

directivity gain in the link from the ℓ-th base station to the typical user, where ϕ_r^ℓ and ϕ_t^ℓ are the angle of arrival and the angle of departure of the signal.

Assumption 2.3.8 (User association)**.** *The typical user is associated with the base station, either LOS or NLOS, that has the smallest path loss $L(R_\ell)$. The serving base station is denoted as X_0. Both the mobile station and its serving base station will estimate channels including angles of arrivals and fading, and then adjust their antenna steering orientations accordingly to exploit the maximum directivity gain. Errors in channel estimation are neglected, and so are errors in time and carrier frequency synchronizations in our work. Thus, the directivity gain for the desired signal link is $D_0 = M_\mathrm{r}M_\mathrm{t}$. For the ℓ-th interfering link, the angles ϕ_r^ℓ and ϕ_t^ℓ are assumed to be independently and uniformly distributed in $(0, 2\pi]$, which gives a random directivity gain D_ℓ.*

By Assumption 2.3.7 and Assumption 2.3.8, the directivity gain in an interference link D_ℓ is a discrete random variable with the probability distribution as $D_\ell = a_k$ with probability b_k ($k \in \{1, 2, 3, 4\}$), where a_k and b_k are constants defined in Table 2.1, $c_\mathrm{r} = \frac{\theta_\mathrm{r}}{2\pi}$, and $c_\mathrm{t} = \frac{\theta_\mathrm{t}}{2\pi}$.

Assumption 2.3.9 (Small-scale fading)**.** *We assume independent Nakagami fading for each link. Different parameters of Nakagami fading N_L and N_N are*

assumed for LOS and NLOS links. Let h_ℓ be the small-scale fading term on the ℓ-th link. Then $|h_\ell|^2$ is a normalized Gamma random variable. Further, for simplicity, we assume N_{L} and N_{N} are positive integers. We also ignore the frequency selectivity in fading, as measurements show that the delay spread is generally small [120], and the impact of frequency-selective fading can be minimized by techniques like orthogonal frequency-division multiplexing or frequency domain equalization [59].

Measurement results indicated that small-scale fading at mmWave is less severe than that in conventional systems when narrow beam antennas are used [120]. Thus, we can use a large Nakagami parameter N_{L} to approximate the small-variance fading as found in the LOS case. Let σ^2 be the thermal noise power normalized by P_{t}. Based on the assumptions thus far, the SINR received by the typical user can be expressed as

$$\mathrm{SINR} = \frac{|h_0|^2 M_{\mathrm{r}} M_{\mathrm{t}} L(R_0)}{\sigma^2 + \sum_{\ell>0: X_\ell \in \Phi} |h_\ell|^2 D_\ell L(R_\ell)}. \tag{2.2}$$

Note that the SINR in (2) is a random variable, due to the randomness in the base station locations R_ℓ, small-scale fading h_ℓ, and the directivity gain D_ℓ. Using the proposed system model, we will evaluate the mmWave SINR and rate coverage in the following section.

2.4 Coverage and Rate Analysis in General Networks

In this section, we analyze the coverage and rate in the proposed model of a general mmWave network. First, we provide some SINR ordering results

regarding different parameters of the antenna pattern. Then we derive expressions for the SINR and rate coverage probability in mmWave networks with general LOS probability function $p(R)$. To simplify subsequent analysis, we then introduce a systematic approach to approximate $p(R)$ by a moment matched equivalent step function.

2.4.1 Stochastic Ordering of SINR With Different Antenna Geometries

One differentiating feature of mmWave cellular networks is the deployment of directional antenna arrays. Consequently, the performance of mmWave networks will depend on the adaptive array pattern through the beamwidth, the directivity gain, and the back lobe gain. In this section, we establish some results on stochastic ordering of the SINRs in the systems with different antenna geometries. While we will focus on the array geometry at the transmitter, the same results, however, also apply to the receiver array geometry. The concept of stochastic ordering has been applied in analysis of wireless systems [134, 51]. Mathematically, the ordering of random variables can be defined as follows [134, 51].

Definition 2.4.1. Let X and Y be two random variables. X stochastically dominates Y, i.e., X has a better distribution than Y, if $\mathbb{P}(X > t) > \mathbb{P}(Y > t)$ for all $t \in \mathbb{R}$.

Next, define the front-to-back ratio (FBR) at the transmitter ξ_{t} as the ratio between the main lobe directivity gain M_{t} and the back lobe gain m_{t},

i.e., $\xi_{\mathrm{t}} = M_{\mathrm{t}}/m_{\mathrm{t}}$. We introduce the key result on stochastic ordering of the SINR with respect to the directivity gains as follows.

Proposition 2.4.1 (Stochastic ordering w.r.t. directivity gains)**.** *Given a fixed beamwidth θ_{t} and FBR ξ_{t} at the transmitter, the mmWave network with the larger main lobe directivity gain M_{t} has a better SINR distribution. Similarly, with fixed beamwidth θ_{t} and main lobe gain M_{t}, a larger FBR ξ_{t} provides a better SINR distribution.*

Proof. From Definition 2.4.1, we need to show that for each realization of base station locations R_ℓ, small-scale fading h_ℓ, and angles ϕ_{r}^{ℓ} and ϕ_{t}^{ℓ}, the value of the SINR increases with M_{t} and ξ_{t}. Given R_ℓ, h_ℓ, ϕ_{r}^{ℓ}, and ϕ_{r}^{ℓ} ($\ell \in \mathbb{N}$), we can normalize both the numerator and denominator of (2.2) by M_t, and then write $\mathrm{SINR} = \frac{|h_0|^2 M_{\mathrm{r}} L(R_0)}{\sigma^2/M_{\mathrm{t}} + \sum_{\ell>0: X_\ell \in \Phi} \bar{D}_\ell(\xi_{\mathrm{t}}) |h_\ell|^2 L(R_\ell)}$, where $\bar{D}_\ell(\xi_{\mathrm{t}}) = e_k$ with probability b_k, and b_k, e_k are constants defined in Table 2.1. Note that $\bar{D}_\ell(\xi_{\mathrm{t}})$ is independent of M_{t}, and is a non-increasing function of ξ_{t}. Hence, when ξ_{t} is fixed, larger M_{t} provides larger SINR; when M_{t} is fixed, larger ξ_{t} provides larger SINR. □

Next, we provide the stochastic ordering result regarding beamwidth as follows.

Proposition 2.4.2 (Stochastic ordering w.r.t. beamwidth)**.** *Given a fixed main lobe gain M_{t} and FBR ξ_{t} at the transmitter, a smaller beamwidth θ_{t} provides a better SINR distribution.*

The proposition can be rigorously proved using coupling techniques. We omit the proof here and instead provide an intuitive explanation as below. Intuitively, with narrower main lobes, fewer base stations will transmit interference to the typical user via their main lobes, which gives a smaller interference power. The desired signal term in (2.2) is independent of the beamwidth, as we ignore the channel estimation errors and potential angle spread. Hence, based on our model assumptions, smaller beamwidths provide a better SINR performance.

We note that the ordering result in Proposition 2.4.2 assumes that there is no angle spread in the channel. With angle spread, a narrow-beam antenna may capture only the signal energy arriving inside its main lobe, missing the energy spread outside, which causes a gain reduction in the signal power [60]. Consequently, the results in Proposition 2.4.2 should be interpreted as applying to the case where beamwidths are larger than the angle spread, e.g. if the beamwidth is more than 55° per the measurements in [113]. We defer more detailed treatment of angle spread to future work.

2.4.2 SINR Coverage Analysis

The SINR coverage probability $P_{\mathrm{c}}(T)$ is defined as the probability that the received SINR is larger than some threshold $T > 0$, i.e., $P_{\mathrm{c}}(T) = \mathbb{P}(\mathrm{SINR} > T)$. We present the following lemmas before introducing the main results on SINR coverage. By Assumption 2.3.4, the outdoor base station process Φ can be divided into two independent non-homogeneous PPPs: the LOS base

station process Φ_{L} and NLOS process Φ_{N}. We will equivalently consider Φ_{L} and Φ_{N} as two independent tiers of base stations. As the user is assumed to connect to the base station with the smallest path loss, the serving base station can only be either the nearest base station in Φ_{L} or the nearest one in Φ_{N}. The following lemma provides the distribution of the distance to the nearest base station in Φ_{L} and Φ_{N}.

Lemma 2.4.1. *Given the typical user observes at least one LOS base station, the conditional probability density function of its distance to the nearest LOS base station is*

$$f_{\mathrm{L}}(x) = 2\pi\lambda x p(x) \mathrm{e}^{-2\pi\lambda \int_0^x r p(r) \mathrm{d}r} / B_{\mathrm{L}}, \tag{2.3}$$

where $x > 0$, $B_{\mathrm{L}} = 1 - \mathrm{e}^{-2\pi\lambda \int_0^\infty r p(r) \mathrm{d}r}$ is the probability that a user has at least one LOS base station, and $p(r)$ is the LOS probability function defined in Section 2.3. Similarly, given the user observes at least one NLOS base station, the conditional probability density function of the distance to the nearest NLOS base station is

$$f_{\mathrm{N}}(x) = 2\pi\lambda x (1 - p(x)) \mathrm{e}^{-2\pi\lambda \int_0^x r(1-p(r)) \mathrm{d}r} / B_{\mathrm{N}}, \tag{2.4}$$

where $x > 0$, and $B_{\mathrm{N}} = 1 - \mathrm{e}^{-2\pi\lambda \int_0^\infty r(1-p(r)) \mathrm{d}r}$ is the probability that a user has at least one NLOS base station.

Proof. The proof follows [30, Theorem 10] and is omitted here. □

Next, we compute the probability that the typical user is associated with either a LOS or a NLOS base station.

Lemma 2.4.2. *The probability that the user is associated with a LOS base station is*

$$A_{\mathrm{L}} = B_{\mathrm{L}} \int_0^{\infty} \mathrm{e}^{-2\pi\lambda \int_0^{\psi_{\mathrm{L}}(x)} (1-p(t))t\mathrm{d}t} f_{\mathrm{L}}(x)\mathrm{d}x, \tag{2.5}$$

where $\psi_{\mathrm{L}}(x) = (C_{\mathrm{N}}/C_{\mathrm{L}})^{1/\alpha_{\mathrm{N}}}\, x^{\alpha_{\mathrm{L}}/\alpha_{\mathrm{N}}}$. *The probability that the user is associated with a NLOS base station is* $A_{\mathrm{N}} = 1 - A_{\mathrm{L}}$.

Proof. See Section 2.8.2. □

Further, conditioning on that the serving base station is LOS (or NLOS), the distance from the user to its serving base station follows the distribution given in the following lemma.

Lemma 2.4.3. *Given that a user is associated with a LOS base station, the probability density function of the distance to its serving base station is*

$$\hat{f}_{\mathrm{L}}(x) = \frac{B_{\mathrm{L}} f_{\mathrm{L}}(x)}{A_{\mathrm{L}}} \mathrm{e}^{-2\pi\lambda \int_0^{\psi_{\mathrm{L}}(x)} (1-p(t))t\mathrm{d}t}, \tag{2.6}$$

when $x > 0$. *Given the user is served by a NLOS base station, the probability density function of the distance to its serving base station is*

$$\hat{f}_{\mathrm{N}}(x) = \frac{B_{\mathrm{N}} f_{\mathrm{N}}(x)}{A_{\mathrm{N}}} \mathrm{e}^{-2\pi\lambda \int_0^{\psi_{\mathrm{N}}(x)} p(t)t\mathrm{d}t}, \tag{2.7}$$

where $x > 0$, *and* $\psi_{\mathrm{N}}(x) = (C_{\mathrm{L}}/C_{\mathrm{N}})^{1/\alpha_{\mathrm{L}}}\, x^{\alpha_{\mathrm{N}}/\alpha_{\mathrm{L}}}$.

Proof. The proof follows a similar method as that of Lemma 2.4.2, and is omitted here. □

Now, based on Lemma 2.4.2 and Lemma 2.4.3, we present the main theorem on the SINR coverage probability as follows

Theorem 2.4.1. *The SINR coverage probability $P_c(T)$ can be computed as*

$$P_{\mathrm{c}}(T) = A_{\mathrm{L}} P_{\mathrm{c,L}}(T) + A_{\mathrm{N}} P_{\mathrm{c,N}}(T), \tag{2.8}$$

where for $\mathrm{s} \in \{\mathrm{L}, \mathrm{N}\}$, $P_{\mathrm{c},s}(T)$ *is the conditional coverage probability given that the user is associated with a base station in* Φ_{s}. *Further,* $P_{\mathrm{c,s}}(T)$ *can be evaluated as*

$$\begin{aligned} P_{c,\mathrm{L}}(T) \approx & \sum_{n=1}^{N_{\mathrm{L}}} (-1)^{n+1} \binom{N_{\mathrm{L}}}{n} \\ & \times \int_0^{\infty} \mathrm{e}^{-\frac{n\eta_{\mathrm{L}} x^{\alpha_{\mathrm{L}}} T\sigma^2}{C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}}} - Q_n(T,x) - V_n(T,x)} \hat{f}_{\mathrm{L}}(x) \mathrm{d}x, \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} P_{c,\mathrm{N}}(T) \approx & \sum_{n=1}^{N_{\mathrm{N}}} (-1)^{n+1} \binom{N_{\mathrm{N}}}{n} \\ & \times \int_0^{\infty} \mathrm{e}^{-\frac{n\eta_{\mathrm{N}} x^{\alpha_{\mathrm{N}}} T\sigma^2}{C_{\mathrm{N}} M_{\mathrm{r}} M_{\mathrm{t}}} - W_n(T,x) - Z_n(T,x)} \hat{f}_{\mathrm{N}}(x) \mathrm{d}x. \end{aligned} \tag{2.10}$$

where

$$Q_n(T,x) = 2\pi\lambda \sum_{k=1}^{4} b_k \int_x^{\infty} F\left(N_{\mathrm{L}}, \frac{n\eta_{\mathrm{L}} \bar{a}_k T x^{\alpha_{\mathrm{L}}}}{N_{\mathrm{L}} t^{\alpha_{\mathrm{L}}}}\right) p(t) t \mathrm{d}t, \tag{2.11}$$

$$\begin{aligned} V_n(T,x) = & 2\pi\lambda \sum_{k=1}^{4} b_k \int_{\psi_{\mathrm{L}}(x)}^{\infty} F\left(N_{\mathrm{N}}, \frac{nC_{\mathrm{N}}\eta_{\mathrm{L}} \bar{a}_k T x^{\alpha_{\mathrm{L}}}}{C_{\mathrm{L}} N_{\mathrm{N}} t^{\alpha_{\mathrm{N}}}}\right) \\ & (1 - p(t)) t \mathrm{d}t, \end{aligned} \tag{2.12}$$

$$W_n(T,x) = 2\pi\lambda \sum_{k=1}^{4} b_k \int_{\psi_{\mathrm{N}}(x)}^{\infty} F\left(N_{\mathrm{L}}, \frac{nC_{\mathrm{L}}\eta_{\mathrm{N}}\bar{a}_k T x^{\alpha_{\mathrm{N}}}}{C_{\mathrm{N}} N_{\mathrm{L}} t^{\alpha_{\mathrm{L}}}}\right) p(t) t \mathrm{d}t, \tag{2.13}$$

$$Z_n(T,x) = 2\pi\lambda \sum_{k=1}^{4} b_k \int_{x}^{\infty} F\left(N_{\mathrm{N}}, \frac{n\eta_{\mathrm{N}}\bar{a}_k T x^{\alpha_{\mathrm{N}}}}{N_{\mathrm{N}} t^{\alpha_{\mathrm{N}}}}\right) (1-p(t)) t \mathrm{d}t, \tag{2.14}$$

and $F(N,x) = 1 - 1/(1+x)^N$. *For* $s \in \{\mathrm{L},\mathrm{N}\}$, $\eta_s = N_s(N_s!)^{-\frac{1}{N_s}}$, N_s *are the parameters of the Nakagami small-scale fading; for* $k \in \{1,2,3,4\}$, $\bar{a}_k = \frac{a_k}{M_{\mathrm{t}} M_{\mathrm{r}}}$, a_k *and* b_k *are constants defined in Table 2.1.*

Proof. See Section 2.8.3. □

Though as an approximation of the SINR coverage probability, we find that the expressions in Theorem 2.4.1 compare favorably with the simulations in Section 2.6.1. In addition, the expressions in Theorem 2.4.1 compute much more efficiently than prior results in [26], which required a numerical inverse of a Fourier transform. Last, the LOS probability function $p(t)$ may itself have a very complicated form, e.g. the empirical function for small cell simulations in [1], which will make the numerical evaluation difficult. Hence, we propose simplifying the system model by using a step function to approximate $p(t)$ in Section 2.4.4. Before that, we introduce our rate analysis results in the following section.

2.4.3 Rate Analysis

In this section, we analyze the distribution of the achievable rate Γ in mmWave networks. We use the following definition for the achievable rate

$$\Gamma = W \log_2(1 + \min\{\mathrm{SINR}, T_{\max}\}), \tag{2.15}$$

where W is the bandwidth assigned to the typical user, and $T_{\max}$ is a SINR threshold determined by the order of the constellation and the limiting distortions from the RF circuit. The use of a distortion threshold $T_{\max}$ is needed because of the potential for very high SINRs in mmWave that may not be exploited due to other limiting factors like linearity in the radio frequency front-end.

The average achievable rate $\mathbb{E}[\Gamma]$ can be computed using the following Lemma from the SINR coverage probability $P_{\mathrm{c}}(T)$.

Lemma 2.4.4. *Given the SINR coverage probability $P_{\mathrm{c}}(T)$, the average achievable rate in the network is $\mathbb{E}[\Gamma] = \frac{W}{\ln 2} \int_0^{T_{\max}} \frac{P_{\mathrm{c}}(T)}{1+T} \mathrm{d}T$.*

Proof. See [18, Theorem 3] and [7, Section V]. □

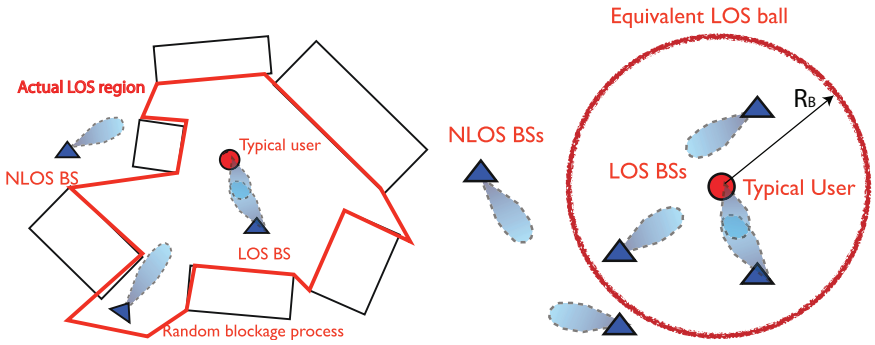
Lemma 2.4.4 provides a first order characterization of the rate distribution. We can also derive the exact rate distribution using the *rate coverage probability* $P_{\mathrm{R}}(\gamma)$, which is the probability that the achievable rate of the typical user is larger than some threshold γ: $P_{\mathrm{R}}(\gamma) = \mathbb{P}[\Gamma > \gamma]$. The rate coverage probability $P_{\mathrm{R}}(\gamma)$ can be evaluated through a change of variables as in the following lemma.

Lemma 2.4.5. *Given the SINR coverage probability $P_{\mathrm{c}}(T)$, for $\gamma < W \log_{\mathrm{N}}(1+T_{\max})$, the rate coverage probability can be computed as $P_{\mathrm{R}}(\gamma) = P_{\mathrm{c}}(2^{\gamma/W}-1)$.*

Proof. The proof is similar to that of [129, Theorem 1]. For $\gamma < W \log_{\mathrm{N}}(1+T_{\max})$, it directly follows that $P_{\mathrm{R}}(\gamma) = \mathbb{P}\left[\mathrm{SINR} > 2^{\gamma/W}-1\right] = P_{\mathrm{c}}(2^{\gamma/W}-1)$. □

Lemma 2.4.5 will allow comparisons to be made between mmWave and conventional systems that use different bandwidths, as presented in Section 2.6.1.

2.4.4 Simplification of LOS Probability Function



(a) Irregular shape of an acutal LOS region.

(b) Approximation using the equivalent LOS ball.

Figure 2.2: Simplification of the random LOS region as a fixed equivalent LOS ball. In (a), we illustrate one realization of randomly located buildings corresponding to a general LOS function $p(x)$. The LOS region observed by the typical user has an irregular shape. In (b), we approximate $p(x)$ by a step function. Equivalently, the LOS region is also approximated by a fixed ball. Only base stations inside the ball are considered LOS to the user.

The expressions in Theorem 2.4.1 generally require numerical evaluation of multiple integrals, and may become difficult to analyze. In this section, we propose to simplify the analysis by approximating a general LOS probability function $p(t)$ by a step function. We denote the step function as $S_{R_{\mathrm{B}}}(x)$, where $S_{R_{\mathrm{B}}}(x) = 1$ when $0 < x < R_{\mathrm{B}}$, and $S_{R_{\mathrm{B}}}(x) = 0$ otherwise. Essentially, the LOS probability of the link is taken to be one within a certain fixed radius R_{B} and zero outside the radius. An interpretation of the simplification is that the irregular geometry of the LOS region in Fig. 2.2 (a) is replaced with its equivalent LOS ball in Fig. 2.2 (b). Such simplification not only provides efficient expressions to compute SINR, but enables simpler analysis of the network performance when the network is dense.

We will propose two criterions to determine the R_{B} given LOS probability function $p(t)$. Before that, we first review some useful facts.

Theorem 2.4.2. *Given the LOS probability function $p(x)$, the average number of LOS base stations that a typical user observes is $\rho = 2\pi\lambda \int_0^\infty p(t)t\mathrm{d}t$.*

Proof. The average number of LOS base stations can be computed as

$$\rho = \mathbb{E}\left[\sum_{X_\ell \in \Phi} \mathbb{I}(X_\ell \in \Phi_{\mathrm{L}})\right] \overset{(a)}{=} 2\pi\lambda \int_0^\infty p(t)t\mathrm{d}t,$$

where (a) follows directly from Campbell's formula of PPP [24]. □

A direct corollary of Theorem 2.4.2 follows as below.

Corollary 2.4.2.1. *When $p(x) = S_R(x)$, the average number of LOS base stations is $\rho = \pi\lambda R^2$.*

Note that Theorem 2.4.2 also indicates that a typical user will observe a finite number of LOS base stations almost surely when $\int_0^\infty p(t)t\mathrm{d}t < \infty$. Hence, if $p(x)$ satisfies $\int_0^\infty p(t)t\mathrm{d}t < \infty$, the parameter R_B in $S_{R_\mathrm{B}}(x)$ can be determined by matching the average number of LOS base stations a user may observe.

Criterion 2.4.1 (Mean LOS BS Number)**.** *When $\int_0^\infty p(t)t\mathrm{d}t < \infty$, the parameter R_B of the equivalent step function $S_{R_\mathrm{B}}(x)$ is determined to match the first moment of ρ. By Theorem 2.4.2, it follows that $R_\mathrm{B} = \left(2\int_0^\infty p(t)t\mathrm{d}t\right)^{0.5}$.*

In the case where $\int_0^\infty p(t)t\mathrm{d}t < \infty$ is not satisfied, another criterion to determine R_B is needed. Note that even if the first moment is infinite, the probability that the user is associated with a LOS base station exists and is naturally finite for all $p(t)$. Hence, we propose the second criterion regarding the LOS association probability as follows.

Criterion 2.4.2 (LOS Association Probability)**.** *Given a LOS probability function $p(t)$, the parameter R_B of its equivalent step function $S_{R_\mathrm{B}}(x)$ is determined such that the LOS association probability A_L is unchanged after approximation.*

From Lemma 2.4.2, the LOS association probability for a step function $S_{R_\mathrm{B}}(x)$ equals $1-\mathrm{e}^{-\lambda\pi R_\mathrm{B}^2}$. Hence, by Criterion 2.4.2, R_B can be determined as $R_\mathrm{B} = \left(\frac{-\ln(1-A_\mathrm{L})}{\lambda\pi}\right)^{0.5}$.

Last, we explain the physical meaning of the step function approximation as follows. As shown in Fig. 2.2(a), with a general LOS probability

function $p(x)$, the buildings are randomly located, and thus the actual LOS region observed by the typical user may have an unusual shape. Although it is possible to incorporate such randomness of the size and shape by integrating over $p(t)$, the expressions with multiple integrals can make the analysis and numerical evaluation difficult [26]. In Fig. 2.2(b), by approximating the LOS probability function as a step function $S_{R_\mathrm{B}}(x)$, we equivalently approximate the LOS region by a fixed ball $\mathcal{B}(0, R_\mathrm{B})$, which we define as the *equivalent LOS ball*. As will be shown in Section 2.5, approximating $p(x)$ as a step function enables fast numerical computation, simplifies the analysis, and provides design insights for dense network. Besides, we will show in simulations in Section 2.6.1 that the error due to such approximation is generally small in dense mmWave networks, which also motivates us to use this first-order approximation of the LOS probability function to simplify the dense network analysis in the following section.

2.5 Analysis of Dense mmWave Networks

In this section we specialize our results to dense networks. This approach is motivated by subsequent numerical results in Section 2.6.1 that show mmWave deployments will be dense if they are expected to achieve significant coverage. We derive simplified expressions for the SINR and provide further insights into system performance in this important asymptotic regime.

2.5.1 Dense Network Model

In this section, we build the dense network model by modifying the system model in Section 2.3 with a few additional assumptions. We say that a mmWave cellular network is *dense* if the average number of LOS base stations observed by the typical user ρ is larger than K, or if its LOS association probability A_{L} is larger than $1-\epsilon$, where K and ϵ are pre-defined positive thresholds. In this chapter, for illustration purpose, we will let $K=1$ and $\epsilon = 5\%$. Further, we say that a network is *ultra-dense* when $\rho > 10$. Note that ρ also equals the relative base station density normalized by the average LOS area, in this special case, as we will explain below.

Now we make some additional assumptions that will allow us to further simplify the network model.

Assumption 2.5.1 (LOS equivalent ball)**.** *The LOS region of the typical user is approximated by its equivalent LOS ball $\mathcal{B}(0, R_{\mathrm{B}})$ as defined in Section 2.4.4.*

By Assumption 2.5.1, the LOS probability function $p(t)$ is approximated by its equivalent step function $S_{R_{\mathrm{B}}}(x)$, and the LOS base station process Φ_{L} is made up of the outdoor base stations that are located inside the LOS ball $\mathcal{B}(0, R_{\mathrm{B}})$. Noting that the outdoor base station process Φ is a homogeneous PPP with density λ, the average number of LOS base stations is $\rho = \lambda \pi R_{\mathrm{B}}^2$, which is the outdoor base station density times the area of the LOS region. For ease of illustration, we call ρ the *the relative density* of a mmWave network. The relative density ρ is equivalently: (i) the average number of LOS

base stations that a user will observe, (ii) the ratio of the average LOS area πR_{B}^2 to the size of a typical cell $1/\lambda$ [24], and (iii) the normalized base station density by the size of the LOS ball. We will show in the next section that the SINR coverage in dense networks is largely determined by the relative density ρ.

Assumption 2.5.2 (No NLOS and noise)**.** *Both NLOS base stations and thermal noise are ignored in the analysis since in the dense regime, the performance is limited by other LOS interferers.*

We show later in the simulations that ignoring NLOS base stations and the thermal noise introduces a negligible error in the performance evaluation.

Assumption 2.5.3 (No Small-scale fading)**.** *Small-scale fading is ignored in the dense network analysis, as the signal power from a nearby mmWave LOS transmitter is found to be almost deterministic in measurements [120].*

Based on the dense network model, the signal-to-interference ratio (SIR) can be expressed as

$$\mathrm{SIR} = \frac{M_{\mathrm{t}} M_{\mathrm{r}} R_0^{-\alpha_{\mathrm{L}}}}{\sum_{\ell: X_\ell \in \Phi \cap \mathcal{B}(0, R_{\mathrm{B}})} D_\ell R_\ell^{-\alpha_{\mathrm{L}}}}. \tag{2.16}$$

Now we compute the SIR distribution in the dense network model.

2.5.2 Coverage Analysis in Dense Networks

Now we present an approximation of the SINR distribution in a mmWave dense network. Our main result is summarized in the following theorem.

Theorem 2.5.1. *The SINR coverage probability in a dense network can be approximated as*

$$P_{\mathrm{c}}(T) \approx \rho \mathrm{e}^{-\rho} \sum_{\ell=1}^{N} (-1)^{\ell+1} \binom{N}{\ell} \int_0^1 \prod_{k=1}^{4} \exp\left(-\frac{2}{\alpha_{\mathrm{L}}} b_k \rho t \right. \\ \left. \times (\ell \eta T \bar{a}_k)^{\frac{2}{\alpha_{\mathrm{L}}}} \Gamma\left(-\frac{2}{\alpha_{\mathrm{L}}}; \ell \eta T \bar{a}_k, \ell \eta T \bar{a}_k t^{\frac{\alpha_{\mathrm{L}}}{2}} \right) \right) \mathrm{d}t, \tag{2.17}$$

where $\Gamma(s; a, b) = \int_a^b x^{s-1} \mathrm{e}^{-x} \mathrm{d}x$ *is the incomplete gamma function,* $\bar{a}_k = a_k/(M_{\mathrm{t}} M_{\mathrm{r}})$*,* a_k *and* b_k *are defined in Table 2.1,* $\eta = N(N!)^{\frac{1}{N}}$*, and* N *is the number of terms used in the approximation.*

Proof. See 2.8.4. □

When $\alpha_{\mathrm{L}} = 2$, the expression in Theorem 2.5.1 can be further simplified as follows.

Corollary 2.5.1.1. *When* $\alpha_{\mathrm{L}} = 2$*, the SINR coverage probability approximately equals*

$$P_{\mathrm{c}}(T) \approx \rho \mathrm{e}^{-\rho} \sum_{\ell=1}^{N} (-1)^{\ell+1} \binom{N}{\ell} \int_0^1 \prod_{k=1}^{4} \exp\left(\rho b_k \times \right. \\ \left. \left(\mathrm{e}^{-\ell \eta T \bar{a}_k t} - t \mathrm{e}^{-\ell \eta T \bar{a}_k} \right) \right) \left(\frac{1 - \mathrm{e}^{-\ell \mu \eta T \bar{a}_k t}}{1 - \mathrm{e}^{-\ell \mu \eta T \bar{a}_k}} \right)^{\ell \eta T b_k \bar{a}_k \rho t} \mathrm{d}t, \tag{2.18}$$

where $\mu = \mathrm{e}^{0.577}$*.*

The results in Theorem 2.5.1 generally provide a close approximation of the SINR distribution when enough terms are used, e.g. when $N \geq 5$, as will be shown in Section 2.6.2. More importantly, we note that the expressions in

Theorem 2.5.2 are very efficient to compute, as most numerical tools support fast evaluation of the gamma function in (2.17), and (2.18) only requires a simple integral over a finite interval. Besides, given the path loss exponent α_{L} and the antenna geometry a_k, b_k, Theorem 2.5.1 shows that the approximated SINR is only a function of the relative density ρ, which indicates the SIR distribution in a dense network is mostly determined on the average number of LOS base station to a user.

2.5.3 Asymptotic Analysis in Ultra-Dense Networks

To obtain further insights into coverage in dense networks, we provide results on the asymptotic SIR distribution when the relative density ρ becomes large. We use this distribution to answer the following questions: (i) What is the asymptotic SIR distribution when the network becomes extremely dense? (ii) Does increasing base station density always improve SIR in a mmWave network?

First, we present the main asymptotic results as follows.

Theorem 2.5.2. *In a dense network, when the LOS path loss exponent $\alpha_{\mathrm{L}} \leq 2$, the SIR converges to zero in probability, as $\rho \to \infty$. When $\alpha_{\mathrm{L}} > 2$, the SIR converges to a nonzero random variable SIR_0 in distribution, as $\rho \to \infty$; Based on [41, Proposition 10], a lower bound of the coverage probability for the asymptotic SIR_0 is that for $T > 1$,*

$$\mathbb{P}(SIR_0 > T) \geq \frac{\alpha_{\mathrm{L}} T^{-2/\alpha_{\mathrm{L}}}}{2\pi \sin(2\pi/\alpha_{\mathrm{L}})}.$$

Proof. The proof is available in a longer version online [28, Appendix E]. □

Note that Theorem 2.5.2 indicates that increasing base station density above some threshold will hurt the system performance, and that the SINR optimal base station density is finite.

Now we provide an intuitive explanation of the asymptotic results as follows. When increasing the base station density, the distances between the user and base stations become smaller, and the user becomes more likely to be associated with a LOS base station. When the density is very high, however, a user sees several LOS base stations and thus experiences significant interference.

We note that the asymptotic trends in Theorem 2.5.2 are valid when base stations are all assumed to be active in the network. A simple way to avoid "over-densification" is to simply turn off a fraction of the base stations. This is a simple kind of interference management; study of more advanced interference management concepts is an interesting topic for future work.

2.6 Numerical Simulations

In this section, we first present some numerical results based on our analyses in Section 2.4 and Section 2.5. We conclude with some simulations using real building distributions to validate our proposed mmWave network model.

2.6.1 General Network Simulations

In this section, we provide numerical simulations to validate our analytical results in Section 2.4, and further discuss their implications on system design. We assume the mmWave network is operated at 28 GHz, and the bandwidth assigned to each user is $W = 100$ MHz. The LOS and NLOS path loss exponents are $\alpha_{\mathrm{L}} = 2$ and $\alpha_{\mathrm{N}} = 4$. The parameters of the Nakagami fading are $N_{\mathrm{L}} = 3$ and $N_{\mathrm{N}} = 2$. We assume the LOS probability function is $p(x) = \mathrm{e}^{-\beta x}$, where $1/\beta = 141.4$ meters. For the ease of illustration, we define the notion of *the average cell radius* of a network as follows. Note that if the base station density is λ, the average cell size in the network is $1/\lambda$ [24]. Therefore, the average cell radius r_{c} in a network is defined as the radius of a ball that has the size of an average cell, i.e., $r_{\mathrm{c}} = \sqrt{1/\pi\lambda}$. The average cell radius not only directly relates to the inter-site distance that is used by industry in base station planning, but also equivalently characterizes the base station density in a network; as a large average cell size indicates a low base station density in the network.

First, we compare the SINR coverage probabilities with different transmit antenna parameters in Fig. 2.3 using Monte Carlos simulations. As shown in Fig. 2.3, when the side lobe gain m_{t} is fixed, better SINR performance is achieved by increasing main lobe gain M_{t} and by decreasing the main lobe beamwidth θ_{t}, as indicated by the analysis in Section 2.4.1.

Next, we compare the LOS association probabilities A_{L} with different average cell radii in Fig. 2.4. The results show that the probability that a user

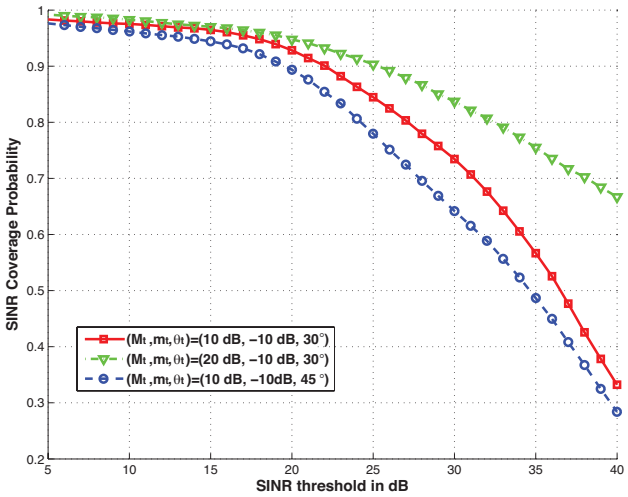


Figure 2.3: SINR coverage probability with different antenna geometry. The average cell radius is $r_{\mathrm{c}} = 100$ meters. The receiver beam pattern is fixed as $G_{10\mathrm{dB},-10\mathrm{dB},90^\circ}$.

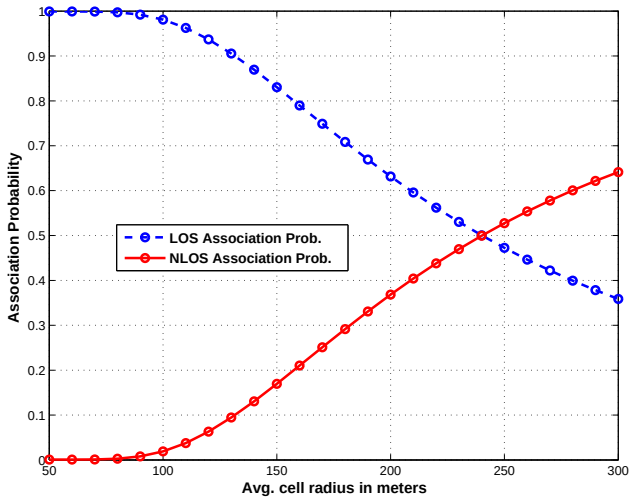
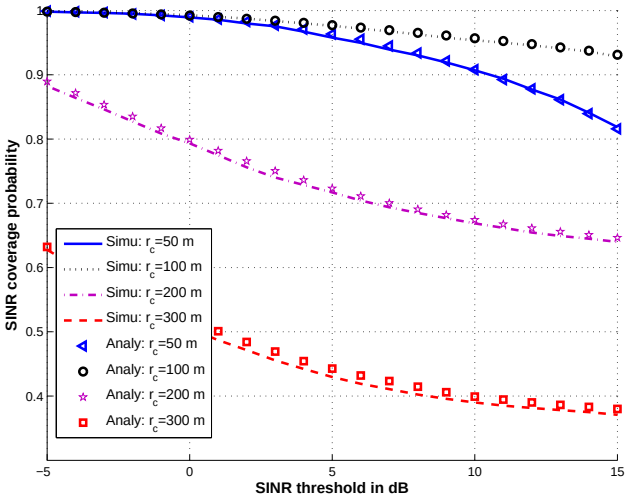
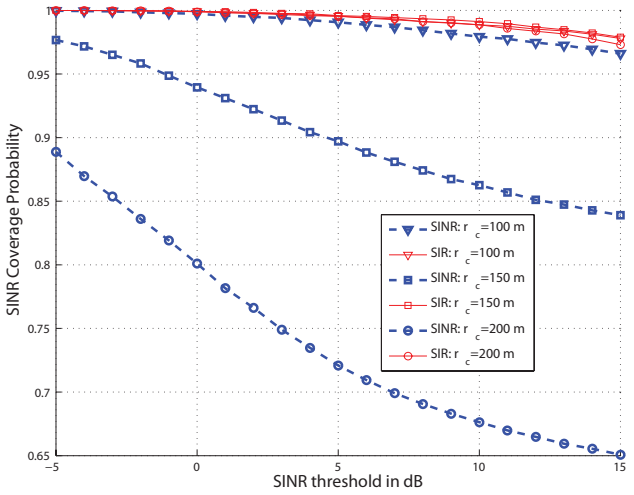


Figure 2.4: LOS association probability with different average cell radii. The lines are drawn from Monte Carlos simulations, and the marks are drawn based on Lemma 2.4.2.



(a) Analytical bounds using Theorem 2.4.1.



(b) Comparison between SINR and SIR.

Figure 2.5: SINR coverage probability with different average cell sizes. The transmit antenna pattern is assumed to be $G_{100\text{dB},0\text{dB},30^\circ}$. In (a), analytical results from Theorem 2.4.1 are shown to provide a tight approximation. In (b), it shows that SIR converges to SINR when the base station density becomes high, which implies that mmWave networks can be interference-limited.

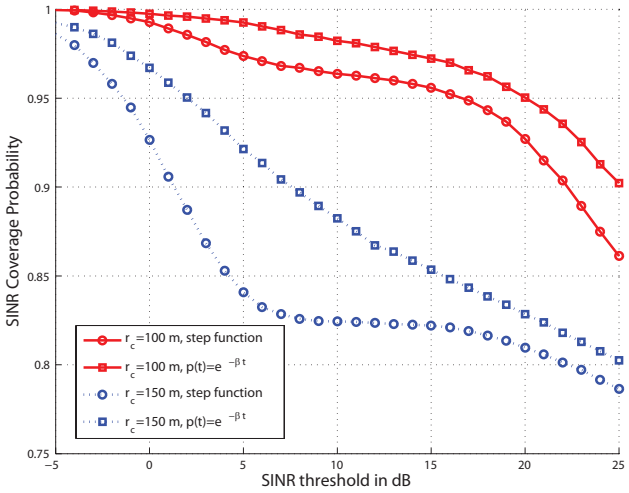


Figure 2.6: Comparison of the SINR coverage between using $p(x)$ and its equivalent step function $S_{R_{\mathrm{B}}}(x)$. The transmit antenna pattern is assumed to be $G_{20\mathrm{dB},-10\mathrm{dB},30^\circ}$. It shows that the step function tends to provide a more pessimistic SINR coverage probability, but the gap becomes smaller as the network becomes more dense.

is associated with a LOS base station increases as the cell radius decreases. The results in Fig. 2.4 also indicate that the received signal power will be mostly determined by the distribution of LOS base stations in a sufficiently dense network, e.g. when the average cell size is smaller than 100 meters in the simulation.

We also compare the SINR coverage probability with different cell radii in Fig. 2.5. The numerical results in Fig. 2.5 (a) show that our analytical results in Theorem 2.4.1 match the simulations well with negligible errors. Unlike in a interference-limited conventional cellular network, where SINR is almost invariant with the base station density [18], the mmWave SINR coverage probability is also shown to be sensitive to the base station density in

Fig. 2.5. The results in Fig. 2.5 (a) also shows that mmWave networks generally require a small cell radius (equivalently a high base station density) to achieve acceptable SINR coverage. Moreover, the results in Fig. 2.5 (b) show that when decreasing average cell radius (i.e., increasing base station density), mmWave networks will transit from power-limited regime into interference-limited regime; as the SIR curves will converge to the SINR curve when densifying the network.

Specifically, comparing the curves for $r_c = 200$ meters and $r_c = 300$ meters in Fig. 2.5 (a), we find that increasing base station density generally improve the SINR in a sparse network; as increasing base station density will increase the LOS association probability and avoid the presence of coverage holes, i.e. the cases that a user observes no LOS base stations. A comparison of the curves for $r_c = 100$ meters and $r_c = 50$ meters, however, also indicates that increasing base station density need not improve SINR, especially when the network is already sufficiently dense. Intuitively, increasing base station density also increases the likelihood to be interfered by strong LOS interferers. In a sufficiently dense network, increasing base station will harm the SINR by adding more strong interferers.

Now we apply Theorem 3 to compare the SINR coverage with different LOS probability functions $p(x)$. We approximate the negative exponential function $p(x) = \mathrm{e}^{-\beta x}$ by its equivalent step function $S_{R_\mathrm{B}}(x)$. Applying either of the criteria in Section 2.4.4, the radius of the equivalent LOS ball R_B equals 200 meters. As shown in Fig. 2.6, the step function approximation generally

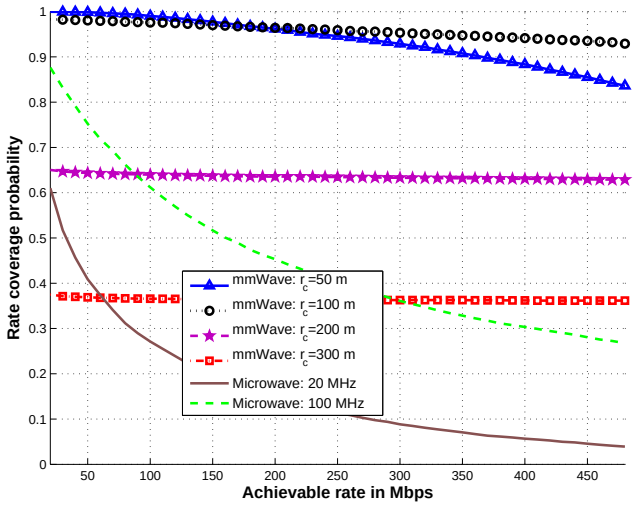


Figure 2.7: Rate coverage comparison between mmWave and conventional cellular networks. The mmWave transmit antenna pattern is assumed to be $G_{10\text{dB},-10\text{dB},30^\circ}$. We assume the conventional system is operated at 2 GHz with a cell radius of 500 m, and the transmit power of the conventional base station is 46 dBm.

provides a lower bound of the actual SINR distribution, and the errors due to the approximation become smaller when the base station density increases. The approximation of step function also enables faster evaluations of the coverage probability, as it simplifies expressions for the numerical integrals.

We provide rate results in Fig. 2.7, where the lines are drawn from Monte Carlos simulations, and the marks are drawn based on Lemma 2.4.5. In the rate simulation, we assume that 64 QAM is the highest constellation supported in the networks, and thus the maximum spectrum efficiency per data stream is 6 bps/Hz. In Fig. 2.7, we compare the rate coverage probability between the mmWave network and a conventional network operated

at 2 GHz. The mmWave bandwidth is 100 MHz (which conceivably could be much larger, e.g. 500 MHz [112, 5]), while we assume the conventional system has a basic bandwidth of 20 MHz, which can be potentially extended to 100 MHz by enabling carrier aggregation [58]. Rayleigh fading is assumed in the UHF network simulations. We further assume that conventional base stations have perfect channel state information, and apply spatial multiplexing (4×4 single user MIMO with zero-forcing precoder) to transmit multiple data streams. More comparison results with other techniques can be found in [25]. Results in Fig. 2.7 shows that, due to the favorable SINR distribution and larger available bandwidth at mmWave frequencies, the mmWave system with a sufficiently small average cell size outperforms the conventional system in terms of providing high data rate coverage.

2.6.2 Dense Network Simulations

Now we show the simulation results based on the dense network analysis in Section 2.5. First, we illustrate the results in Theorem 2.5.1 with the simulations in Fig. 2.8. In the simulations, we include the NLOS base stations and thermal noise, which were ignored in the theoretical derivation. The expression derived in Theorem 2.5.1 generally provides a lower bound of the coverage probability. The approximation becomes more accurate when more terms are used in the approximation, especially when $N \geq 5$. We find that the error due to ignoring NLOS base stations and thermal noise is minor in terms of the SINR coverage probability, primarily impacting low SINRs.

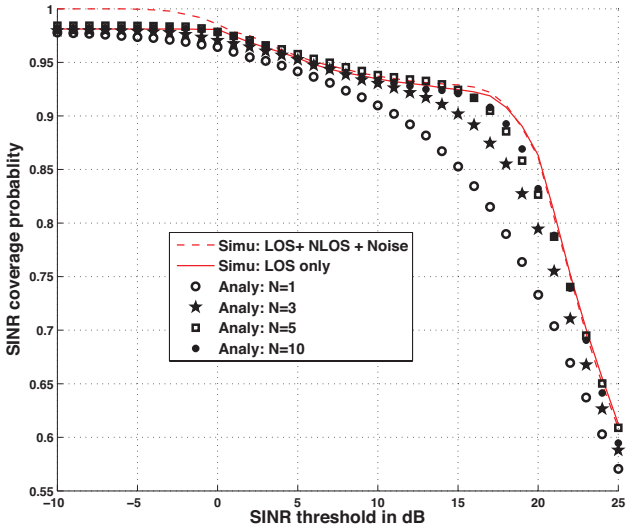


Figure 2.8: Coverage probability in a dense mmWave network. The mmWave transmit antenna pattern is assumed to be $G_{10\text{dB},-10\text{dB},30^\circ}$. We assume $R_{\text{B}} = 200$ m, and the relative base station density $\rho = 4$. N is the number of terms we used to approximate the coverage probability in Theorem 2.5.1.

Next, we compare the SINR coverage probability with different relative base station density when $T = 20$ dB. Recall that $\rho = \lambda\pi R_{\text{B}}^2$ is the base station density normalized by the size of the LOS region. In 2.9(a), the path loss exponent is assumed to be $\alpha = 2$. We compute the coverage probability from $\rho = -20$ dB meters to $\rho = 20$ dB with a step of 1 dB. The analytical expressions in Theorem 2.5.1 are much more efficient than simulations: the plot takes seconds to finish using the analytical expression, while it approximately takes an hour to simulate 10,000 realizations at each step. As shown in Fig. 2.9 (a), although there is some gap between the simulation and the analytical results in the ultra-dense network regime, both curves achieve their maxima at approximately $\rho = 5$, i.e., when the average cell radius r_c is approximately

Table 2.2: Achievable Rate with Different BS Densities

Carrier frequency	28 GHz	28 GHz	28 GHz	2 GHz
Base station density	Ultra dense	Dense	Sparse	-
Relative density ρ	16	4	0.45	-
Spectrum efficiency (bps/Hz)	5.5	5.8	2.7	4.6
Signal bandwith (MHz)	100	100	100	20
Achievable rate (Mpbs)	550	580	270	92

1/2 of the LOS range R_{B}. Moreover, when the base station density grows very large, the coverage probability begins to decrease, which matches the asymptotic results in Theorem 2.5.2. The results also indicate that networks in the environments with dense blockages, e.g. the downtown areas of large cities where the LOS range R_{B} is small, will benefit from network densification; as they are mostly operated in the region where the relative density is (much) smaller than the optimal value $\rho \approx 5$, and thus increasing ρ by densifying networks will improve SINR coverage.

We also simulate with other LOS path loss exponents in Fig. 2.9 (b). The results show that the optimal base station density is generally insensitive to the change of the path loss exponent. When the LOS path loss exponent increases from 1.5 to 2.5, the optimal cell size is almost the same. The results also illustrate that the networks with larger path loss exponent α_{L} have better SINR coverage in the ultra-dense regime when $\rho > 10$. Intuitively, signals attenuate faster with a larger path loss exponent, and thus the inter-cell interference becomes weaker, which motivates a denser deployment of base stations in the network with higher path loss.

Finally, we compare the spectral efficiency and average achievable rates as a function of the relative density ρ in Table 2.2. We find with a reasonable amount of density, e.g. when the relative density ρ is approximately 1, the mmWave system can provide comparable spectrum efficiency as the conventional system at UHF frequencies. With high density, rates that can be achieved are an order of magnitude better than that in the conventional networks, due to the favourable SINR distribution and larger available bandwidth at mmWave frequencies.

2.6.3 Comparison with Real-scenario Simulations

Now we compare our proposed network models with the simulations using real data. In the real-scenario simulations, we use the building distribution on the campus at The University of Texas at Austin. We also apply a modified version of the base station antenna pattern in [3] with a smaller beam width of 30°. The directivity gain at the base station is $M_{\mathrm{t}} = 20$ dB. The mobile station is assumed to use uniform linear array with 4 antennas. When applying our analytical models, we fit the parameters of the LOS probability functions to match the building statistics [30], and use the sectored model for beamforming pattern. We also assume the mmWave base stations are distributed as a PPP with $r_c = 50$ m. As shown in Fig. 2.10, though some deviations in the high SINR regime, our analytical models generally show a close characterization of the reality. The deviation is explained as follows: the proposed analytical model computes the aggregated SINR coverage probabil-

ity, averaging over all realizations of building distributions over the infinite plane, while the real-scenario curve only considers a specific realization of buildings in a finite snapshot window. In this case, our model overestimates the coverage probability in the low SINR regime, and underestimates in the high SINR regime, as both signals and interference become more likely to be blocked in the real scenario simulation. We have found in other simulation examples that the reverse can also be true. Our model should be viewed as a characterization of the average distribution and does not necessarily lower or upper bound the distribution for a given realization.

2.7 Conclusion

In this chapter, I proposed a stochastic geometry framework to analyze coverage and rate in mmWave networks for outdoor users and outdoor infrastructure. Our model took blockage effects into account by applying a distance-dependent LOS probability function, and modeling the base stations as independent inhomogeneous LOS and NLOS point processes. Based on the proposed framework, I derived expressions for the downlink SINR and rate coverage probability in mmWave cellular networks, which were shown to be efficient in computation and also a good fit with the simulations. We further simplified the blockage model by approximating the random LOS region as a fixed-size equivalent LOS ball. Applying the simplified framework, I analyzed the performance and asymptotic trends in dense networks.

We used numerical results to draw several important conclusions about

coverage and rate in mmWave networks.

- SINR coverage can be comparable to conventional networks at UHF frequency when the base station density is sufficiently high.
- Achievable rates can be significantly higher than in conventional networks, thanks to the larger available bandwidth.
- The SINR and rate performance is largely determined by the relative base station density, which is the ratio of the base station density to the blockage density.
- A transition from a power-limited regime to an interference-limited regime is also observed in mmWave networks, when increasing base station density.
- The optimal SINR and rate coverage can be achieved with a finite base station density; as increasing base station density need not improve SINR in a (ultra) dense mmWave network.

2.8 Appendix

2.8.1 Alzer's Lemma

We provide two useful inequalities in the following lemmas. The first lemma approximates the tail probability of a gamma random variable.

Lemma 2.8.1 (From [14])**.** *Let g be a normalized gamma random variable with parameter N. For a constant $\gamma > 0$, the probability $\mathbb{P}(g < \gamma)$ can be*

tightly upper bounded by

$$\mathbb{P}(g < \gamma) < \left[1 - \mathrm{e}^{-a\gamma}\right]^N,$$

where $a = N(N!)^{-\frac{1}{N}}$.

The following inequality will be used in the dense network analysis.

Lemma 2.8.2 (From [14])**.** *For* $x > 0$*, it holds that*

$$-\log(1 - \mathrm{e}^{-ax}) \leq \int_x^\infty \frac{\mathrm{e}^{-t}}{t}\mathrm{d}t \leq -\log(1 - \mathrm{e}^{-bx}),$$

where $a = \mathrm{e}^{0.5772}$ *and* $b = 1$*. Further, the lower bound generally provides a close approximation.*

2.8.2 Proof of Lemma 2.4.2

For $s = \{\mathrm{L}, \mathrm{N}\}$, let d_s be the distance from the typical user to its nearest base station in Φ_s. Note that it is possible that the user observes no base stations in Φ_s. The user is associated with a base station in Φ_L if and only if it has a LOS base station, and its nearest base station in Φ_L has smaller path loss than that of the nearest base station in Φ_N. Hence, it follows that

$$\begin{aligned} A_\mathrm{L} &= B_\mathrm{L}\mathbb{P}\left(C_\mathrm{L} d_\mathrm{L}^{-\alpha_\mathrm{L}} > C_\mathrm{N} d_\mathrm{N}^{-\alpha_\mathrm{N}}\right) \\ &\overset{(a)}{=} B_\mathrm{L} \int_0^\infty \mathbb{P}\left(d_\mathrm{N} > \psi_\mathrm{L}(x)\right) f_\mathrm{L}(x)\mathrm{d}x, \end{aligned} \tag{2.19}$$

where B_L is the probability that the user has at least one LOS base stations, (a) follows that by Lemma 2.4.1, and $f_\mathrm{L}(x)$ is the probability density function of d_L. Next, note that

$$\mathbb{P}\left(d_\mathrm{N} > \psi(x)\right) = \mathbb{P}\left(\Phi_\mathrm{N} \cap \mathcal{B}(0, \psi_\mathrm{L}(x)) = \emptyset\right)$$

$$= \mathrm{e}^{-2\pi\lambda \int_0^{\psi_{\mathrm{L}}(x)} (1-p(t))t\mathrm{d}t}, \tag{2.20}$$

where $\mathcal{B}(0,x)$ denotes the ball centered at the origin of radius x. Substituting (2.20) for (2.19) gives Lemma 2.4.2.

2.8.3 Proof of Theorem 2.4.1

Given that the user is associated with a base station in Φ_{L}, by Slivnyak's Theorem [24], the conditional coverage probability can be computed as

$$P_{\mathrm{c,L}}(T) = \int_0^{\infty} \mathbb{P}\left[h_0 > x^{\alpha_{\mathrm{L}}} T\left(\sigma^2 + I_{\mathrm{L}} + I_{\mathrm{N}}\right)/(C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}})\right] \hat{f}_{\mathrm{L}}(x)\mathrm{d}x, \tag{2.21}$$

where I_{L} and I_{N} defined as

$$I_{\mathrm{L}} = \sum_{\ell: X_\ell \in \Phi_{\mathrm{L}} \cap \bar{\mathcal{B}}(0,x)} C_{\mathrm{L}} |h_\ell|^2 D_\ell R_\ell^{-\alpha_{\mathrm{L}}},$$

$$I_{\mathrm{N}} = \sum_{\ell: X_\ell \in \Phi_{\mathrm{N}} \cap \bar{\mathcal{B}}(0,\psi_{\mathrm{L}}(x))} C_{\mathrm{N}} |h_\ell|^2 D_\ell R_\ell^{-\alpha_{\mathrm{N}}}$$

are the interference strength from the tiers of LOS and NLOS base stations, respectively. Next, noting that $|h_0|^2$ is a normalized gamma random variable with parameter N_{L}, I have the following approximation

$$\begin{aligned}
&\mathbb{P}\left[h_0 > x^{\alpha_{\mathrm{L}}} T\left(\sigma^2 + I_{\mathrm{L}} + I_{\mathrm{N}}\right)/(C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}})\right] \\
&\overset{(a)}{\approx} 1 - \mathbb{E}_{\Phi}\left[\left(1 - \mathrm{e}^{-\frac{\eta_{\mathrm{L}} x^{\alpha_{\mathrm{L}}} T\left(\sigma^2 + I_{\mathrm{L}} + I_{\mathrm{N}}\right)}{C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}}}}\right)^{N_{\mathrm{L}}}\right] \\
&\overset{(b)}{=} \sum_{n=1}^{N_{\mathrm{L}}} (-1)^{n+1} \binom{N_{\mathrm{L}}}{n} \mathbb{E}_{\Phi}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}} x^{\alpha_{\mathrm{L}}} T\left(\sigma^2 + I_{\mathrm{L}} + I_{\mathrm{N}}\right)}{C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}}}}\right] \\
&\overset{(c)}{=} \sum_{n=1}^{N_{\mathrm{L}}} (-1)^{n+1} \binom{N_{\mathrm{L}}}{n} \mathrm{e}^{-\frac{n\eta_{\mathrm{L}} x^{\alpha_{\mathrm{L}}} T\sigma^2}{C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}}}} \mathbb{E}_{\Phi_{\mathrm{L}}}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}} x^{\alpha_{\mathrm{L}}} T I_{\mathrm{L}}}{C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}}}}\right] \mathbb{E}_{\Phi_{\mathrm{N}}}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}} x^{\alpha_{\mathrm{L}}} T I_{\mathrm{N}}}{C_{\mathrm{L}} M_{\mathrm{r}} M_{\mathrm{t}}}}\right],
\end{aligned} \tag{2.22}$$

where $\eta_{\mathrm{L}} = N_{\mathrm{L}}(N_{\mathrm{L}}!)^{-\frac{1}{N_{\mathrm{L}}}}$, (a) is from Lemma 2.8.1 [14] in Section 2.8.1, (b) follows from Binomial theorem and the assumption that N_{L} is an integer, and (c) follows from the fact that Φ_{L} and Φ_{N} are independent. Now I apply concepts from stochastic geometry to compute the term for LOS interfering links $\mathbb{E}_{\Phi_{\mathrm{L}}}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}}x^{\alpha_{\mathrm{L}}}TI_{\mathrm{L}}}{C_{\mathrm{L}}M_{\mathrm{r}}M_{\mathrm{t}}}}\right]$ in (2.22) as

$$\begin{aligned}
&\mathbb{E}_{\Phi_{\mathrm{L}}}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}}x^{\alpha_{\mathrm{L}}}TI_{\mathrm{L}}}{C_{\mathrm{L}}M_{\mathrm{r}}M_{\mathrm{t}}}}\right] \\
&\quad= \mathbb{E}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}}x^{\alpha_{\mathrm{L}}}T\sum_{\ell:X_\ell\in\Phi_{\mathrm{L}}\cap\bar{\mathcal{B}}(0,x)}|h_\ell|^2D_\ell R_\ell^{-\alpha_{\mathrm{L}}}}{M_{\mathrm{r}}M_{\mathrm{t}}}}\right] \\
&\quad\overset{(c)}{=} \mathrm{e}^{\left(-2\pi\lambda\sum_{k=1}^{4}b_k\int_x^\infty\left(1-\mathbb{E}_g\left[\mathrm{e}^{-nT\eta_{\mathrm{L}}g\bar{a}_k(x/t)^{\alpha_{\mathrm{L}}}}\right]\right)p(t)t\mathrm{d}t\right)} \\
&\quad\overset{(d)}{=} \prod_{k=1}^{4}\mathrm{e}^{-2\pi\lambda b_k\int_x^\infty\left(1-1/(1+\eta_{\mathrm{L}}\bar{a}_knT(x/t)^{\alpha_{\mathrm{L}}}/N_{\mathrm{L}})^{N_{\mathrm{L}}}\right)p(t)t\mathrm{d}t} \\
&\quad= \mathrm{e}^{-Q_n(T,x)},
\end{aligned}$$

where g in (c) is a normalized gamma random variable with parameter N_{L}, $\bar{a}_k = \frac{a_k}{M_{\mathrm{t}}M_{\mathrm{r}}}$, and for $1 \le k \le 4$, a_k and b_k are defined previously in Table 2.1; (c) is from computing the Laplace functional of the PPP Φ_{L} [24]; (d) is by computing the moment generating function of a gamma random variable g.

Similarly, for the NLOS interfering links, the small-scale fading term $|h_\ell|^2$ is a normalized gamma variable with parameter N_{N}. Thus, I can compute $\mathbb{E}_{\Phi_{\mathrm{N}}}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}}x^{\alpha_{\mathrm{L}}}TI_{\mathrm{N}}}{C_{\mathrm{L}}M_{\mathrm{r}}M_{\mathrm{t}}}}\right]$ as

$$\begin{aligned}
&\mathbb{E}_{\Phi_{\mathrm{N}}}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}}x^{\alpha_{\mathrm{L}}}TI_{\mathrm{N}}}{C_{\mathrm{L}}M_{\mathrm{r}}M_{\mathrm{t}}}}\right] \\
&\quad= \mathbb{E}\left[\mathrm{e}^{-\frac{n\eta_{\mathrm{L}}x^{\alpha_{\mathrm{L}}}TC_{\mathrm{N}}\sum_{\ell:X_\ell\in\Phi_{\mathrm{N}}\cap\bar{\mathcal{B}}(0,\psi(x))}|h_\ell|^2D_\ell R_\ell^{-\alpha_{\mathrm{N}}}}{C_{\mathrm{L}}M_{\mathrm{r}}M_{\mathrm{t}}}}\right]
\end{aligned}$$

$$= \prod_{k=1}^{4} \mathrm{e}^{-2\pi\lambda b_k \int_{\psi_{\mathrm{L}}(x)}^{\infty} \left(1-1/\left(1+\frac{\eta_{\mathrm{L}}\bar{a}_k nTC_{\mathrm{N}}x^{\alpha_{\mathrm{L}}}}{C_{\mathrm{L}}t^{\alpha_{\mathrm{N}}}N_{\mathrm{N}}}\right)^{N_{\mathrm{N}}}\right)(1-p(t))t\mathrm{d}t}$$
$$= \mathrm{e}^{-V_n(T,x)}.$$

Then, I obtain (2.9) from (2.22) by the linearity of integrals.

Given the user is associated with a NLOS base station, I can also derive the conditional coverage probability $P_{c,\mathrm{N}}(T)$ following same approach as that of $P_{c,\mathrm{L}}(T)$. Thus, I omit the detailed proof of (2.10) here.

Finally, by the law of total probability, it follows that $P_{\mathrm{c}}(T) = A_{\mathrm{L}}P_{\mathrm{c,L}}(T) + A_{\mathrm{N}}P_{\mathrm{c,N}}(T)$.

2.8.4 Proof of Theorem 2.5.1

For a general α_{L}, the coverage probability can be computed as

$$\begin{aligned} P_{\mathrm{c}}(T) &= A_{\mathrm{L}}P_{\mathrm{c,L}}(T) = A_{\mathrm{L}}\mathbb{P}(\mathrm{SIR} > T) \\ &= A_{\mathrm{L}}\int_0^{R_{\mathrm{B}}} \mathbb{P}(C_{\mathrm{L}}M_rM_tr^{-\alpha_{\mathrm{L}}} > TI_r)\frac{2\pi\lambda r}{A_{\mathrm{L}}}\mathrm{e}^{-\lambda\pi r^2}\mathrm{d}r, \end{aligned}$$

where $I_r = \sum_{X_\ell \in \Phi\cap(\mathcal{B}(0,R_{\mathrm{B}})/\mathcal{B}(0,R_{\mathrm{B}}))} D_\ell C_{\mathrm{L}} R_\ell^{-\alpha_{\mathrm{L}}}$ is the interference power given that the distance to the user's serving base station is $R_0 = r$. Next, the probability $\mathbb{P}(C_{\mathrm{L}}M_rM_tr^{-\alpha_{\mathrm{L}}} > TI_r)$ can be approximated as

$$\begin{aligned} &\mathbb{P}(C_{\mathrm{L}}M_rM_tr^{-\alpha_{\mathrm{L}}} > TI_r) \overset{(a)}{\approx} \mathbb{P}(g > Tr^{\alpha_{\mathrm{L}}}I_r/(C_{\mathrm{L}}M_rM_t)) \\ &\overset{(b)}{\approx} 1 - \mathbb{E}_{\Phi_{\mathrm{L}}}\left[\left(1-\mathrm{e}^{-\eta Tr^{\alpha_{\mathrm{L}}}I_r/(C_{\mathrm{L}}M_rM_t)}\right)^N\right] \\ &= \sum_{\ell=1}^{N}\binom{N}{\ell}(-1)^\ell \mathbb{E}_{\Phi_{\mathrm{L}}}\left[\mathrm{e}^{-\ell\eta Tr^{\alpha_{\mathrm{L}}}I_r/(C_{\mathrm{L}}M_rM_t)}\right]. \end{aligned} \tag{2.23}$$

In (a), the dummy variable g is a normalized gamma variable with parameter N, and the approximation in (a) follows from the fact that a normalized Gamma distribution converges to identity when its parameter goes to infinity, i.e., $\lim_{n\to\infty} \frac{n^n x^{n-1}\mathrm{e}^{-nx}}{\Gamma(n)} = \delta(x-1)$ [20], where $\delta(x)$ is the Dirac delta function. In (b), it directly follows from Lemma 2.8.1 by taking $\eta = N(N!)^{1/N}$.

Next, I can compute $\mathbb{E}_{\Phi_{\mathrm{L}}}\left[\mathrm{e}^{-\ell\eta T r^{\alpha_{\mathrm{L}}} I_r/(C_{\mathrm{L}} M_r M_t)}\right]$ as

$$
\begin{aligned}
&\mathbb{E}_{\Phi_{\mathrm{L}}}\left[\mathrm{e}^{-\ell\eta T r^{\alpha_{\mathrm{L}}} I_r/(C_{\mathrm{L}} M_r M_t)}\right] \\
&\overset{(c)}{=} \exp\left(\sum_{k=1}^{4} -2\pi\lambda b_k \int_r^{R_{\mathrm{B}}} 1-\mathrm{e}^{-\ell\eta\bar{a}_k T (r/t)^{\alpha_{\mathrm{L}}}} t\mathrm{d}t\right) \\
&\overset{(d)}{=} \mathrm{e}^{-\pi\lambda(R_{\mathrm{B}}^2 - r^2)} \times \\
&\quad \mathrm{e}^{\sum_{k=1}^{4} \frac{2}{\alpha_{\mathrm{L}}} \pi\lambda r^2 b_k (T\ell\eta\bar{a}_k)^{2/\alpha_{\mathrm{L}}} \int_{\ell\eta T\bar{a}_k (r/R_{\mathrm{B}})^{\alpha_{\mathrm{L}}}}^{\ell\eta T\bar{a}_k} \frac{\mathrm{e}^{-s}}{s^{1+2/\alpha_{\mathrm{L}}}} \mathrm{d}s} \qquad (2.24) \\
&= \mathrm{e}^{-\pi\lambda(R_{\mathrm{B}}^2 - r^2)} \times \\
&\quad \mathrm{e}^{\sum_{k=1}^{4} \frac{2}{\alpha_{\mathrm{L}}} \pi\lambda r^2 b_k (T\ell\eta\bar{a}_k)^{2/\alpha_{\mathrm{L}}} \Gamma\left(\frac{-2}{\alpha_{\mathrm{L}}}; \ell\eta T\bar{a}_k (r/R_{\mathrm{B}})^{\alpha_{\mathrm{L}}}, \ell\eta T\bar{a}_k\right)}, \qquad (2.25)
\end{aligned}
$$

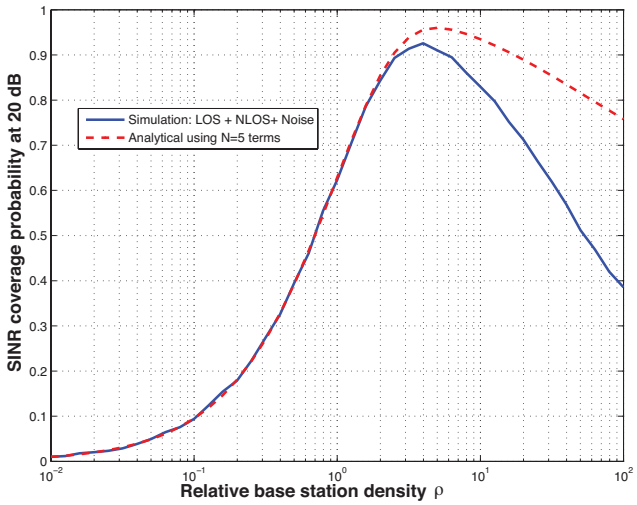
where (c) follows from computing the Laplace functional of the PPP Φ_{L} [24], and (d) follows from changing variable as $s = \ell\eta\bar{a}_k T (r/t)^{\alpha_{\mathrm{L}}}$. Hence, (2.17) directly follows from substituting (2.25) for (2.23) and letting $\rho = \pi\lambda R_{\mathrm{B}}^2$.

When $\alpha_{\mathrm{L}} = 2$, the steps above hold true till (2.24), which can be further simplified as

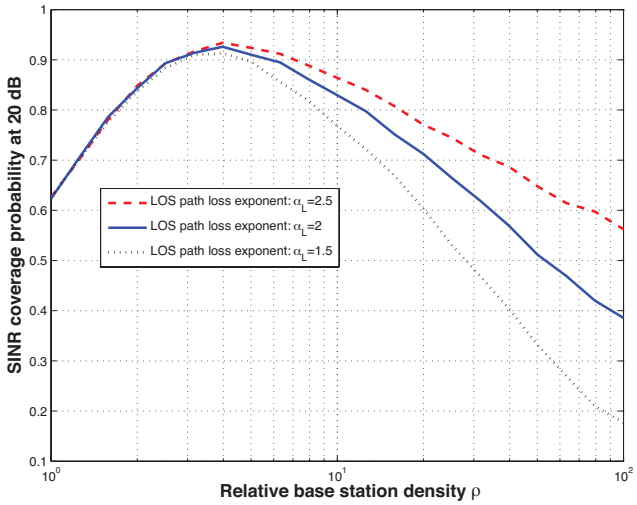
$$
\begin{aligned}
&\mathbb{E}_{\Phi_{\mathrm{L}}}\left[\mathrm{e}^{-\ell\eta T r^{\alpha_{\mathrm{L}}} I_r/(C_{\mathrm{L}} M_r M_t)}\right] \\
&= \mathrm{e}^{-\pi\lambda(R_{\mathrm{B}}^2 - r^2)} \mathrm{e}^{\sum_{k=1}^{4} \left(\pi\lambda r^2 b_k T\ell\eta\bar{a}_k \int_{\ell\eta T\bar{a}_k (r/R_{\mathrm{B}})^2}^{\ell\eta T\bar{a}_k} \frac{\mathrm{e}^{-s}}{s^2} \mathrm{d}s\right)} \\
&\overset{(e)}{=} \mathrm{e}^{-\pi\lambda(R_{\mathrm{B}}^2 - r^2)} \times
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{e}^{\sum_{k=1}^{4}\pi\lambda r^2 b_k\left(\frac{\mathrm{e}^{-(r/R_{\mathrm{B}})^2\ell T\ell\eta\bar{a}_k}}{(r/R_{\mathrm{B}})^2}-\mathrm{e}^{-\ell T\ell\eta\bar{a}_k}+\int_{\ell T\ell\eta\bar{a}_k(r/R_{\mathrm{B}})^2}^{\ell T\ell\eta\bar{a}_k}\frac{\mathrm{e}^{-s}}{s}\mathrm{d}s\right)}\\
&\overset{(f)}{\approx}\mathrm{e}^{-\pi\lambda(R_{\mathrm{B}}^2-r^2)}\exp\left(\sum_{k=1}^{4}\pi\lambda r^2 b_k\left(\frac{\mathrm{e}^{-(r/R_{\mathrm{B}})^2\ell T\ell\eta\bar{a}_k}}{(r/R_{\mathrm{B}})^2}\right.\right.\\
&\left.\left.-\mathrm{e}^{-\ell T\ell\eta\bar{a}_k}-\log\frac{1-\mathrm{e}^{\mu\ell T\ell\eta\bar{a}_k(r/R_{\mathrm{B}})^2}}{1-\mathrm{e}^{\mu\ell T\ell\eta\bar{a}_k}}\right)\right),
\end{aligned}
\tag{2.26}
$$

where (e) is from computing integration by part, (f) follows from Lemma 2.8.2 by letting $\mu = \mathrm{e}^{0.5772}$. Lastly, (2.18) follows from substituting (2.26) for (2.23) and letting $\rho = \pi\lambda R_{\mathrm{B}}^2$.



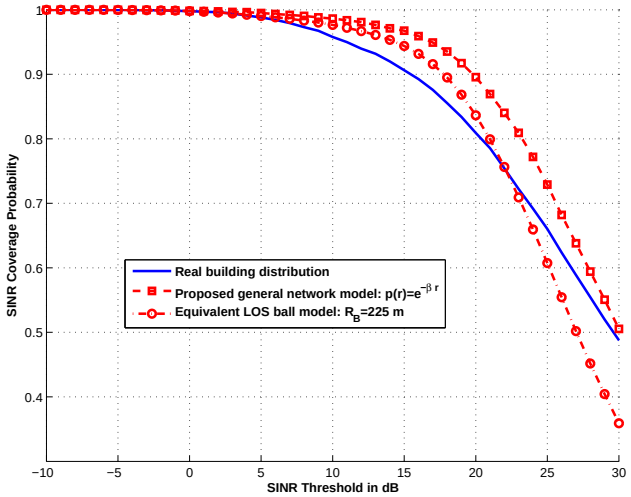
(a) Analytical results using Theorem 2.5.1.



(b) Optimal density with different path loss exponents.

Figure 2.9: SINR coverage probability with different relative base station density when the target SINR=20 dB. In the simulations, we include the NLOS base stations outside the LOS region and the thermal noise. We also fix the radius of the LOS ball as R_{B} =200 meters, and change the base station density λ at each step according to the value of the relative base station density ρ. In (a), it shows that ignoring NLOS base stations and the noise power causes minor errors in terms of the optimal cell radius. In (b), we search for the optimal relative density with different LOS path loss exponents. It shows that the optimal cell radius is generally insensitive to the path loss exponent.

(a) Snapshot of the simulated area from Google map.



(b) Comparison of SINR distribution

Figure 2.10: Comparison of SINR coverage results with real-scenario simulations. The snapshot of The University of Texas at Austin campus is from Google map. We use the actual building distribution of the area in the real-scenario simulation. In the simulations of our proposed analytical models, we let $\beta = 0.0063$ m^{-1} in the LOS probability function $p(r) = e^{-\beta r}$, and $R_{\mathrm{B}} = 225$ m in the simplified equivalent LOS ball model, to match the building statistics in the area [30].

Chapter 3

Analysis of Uplink Massive MIMO Networks

3.1 Motivation and Related Work

Massive MIMO is an approach to increase the area spectrum efficiency in 5G cellular systems [90, 76, 87, 42]. By deploying large-scale antenna arrays, base stations can use multi-user MIMO to serve a large number of users and provide high cell throughput [90, 76, 87, 42]. In this chapter, I focus on the defacto massive MIMO systems operated below 6 GHz, where pilot-aided channel estimation is performed in the uplink, and pilots are reused across cells to reduce the training overhead [90, 76, 87, 42]. Prior work showed that when the number of base station antennas grows large, high throughput is achieved through simple signal processing, and that the asymptotic performance of massive MIMO (in the limit of the number of base station antennas) is limited by pilot contamination [90].

In this chapter, I derive the signal-to-interference ratio (SIR) distribution for the uplink of a massive MIMO network with MRC and ZF receivers, for a random base station topology. The performance with MRC and ZF beamforming in terms of SINR, spectrum efficiency, and energy efficiency was examined in a simple network topology, e.g. in [147, 90, 73, 64, 80, 106],

where the SIR and rate expressions were conditioned on specific user locations or equivalently the received power for each user. The conclusions drawn from the conditional expression, however, need not apply to the spatial average system-level performance due to the difference in users' path losses. For example, the linear scaling between the number of users and antennas examined in [64] does not maintain the uplink SIR distribution, as will be shown in our analysis. This motivates the analysis of the spatial average performance over different base station and user distributions in large-scale massive MIMO networks, which was mainly studied using Monte Carlo simulations in prior work [39, 22].

Stochastic geometry provides a powerful tool to analyze system-level performance in a large-scale network with randomly distributed base stations and users. Assuming a single antenna at each base station, the spatial average downlink SIR and rate distributions were derived for a network with Poisson point process (PPP) distributed base stations, and were shown a reasonable fit with simulations using real base station data [16]. The stochastic geometry framework in [16] was further extended to analyze the performance of MIMO networks: the downlink SIR and rate of multi-user MIMO cellular system were analyzed, e.g. in [51, 45, 145, 77] assuming perfect CSI, and in [72] with quantized CSI from limited feedback. For uplink analysis, prior work [108, 53, 130] showed that the uplink and downlink SIR follows different distributions, due to the difference in network topology. In [130], a stochastic geometry uplink model was proposed to take account the pairwise correlations in the

user locations, where the SIR distributions derived based on the analytical model were shown a good fit with the simulations. The prior results in [16, 51, 45, 72, 145, 77, 108, 53, 130], however, do not directly apply to analyze uplink massive MIMO networks, as (i) they did not take account for the effects of pilot contamination, which becomes a limiting factor with large numbers of antennas [90]; (ii) the analysis in [16, 51, 45, 72, 145, 77] was intended for downlink performance, which follows different distributions from the uplink network; and (iii) the results in [51, 45, 145, 77] were intended for MIMO networks with a few antennas, where the computational complexity for the analytical expressions grows with the number of antennas, and hinders the direct application to the massive MIMO scenarios.

Stochastic geometry was also applied to study the asymptotic SIR and rate in a massive MIMO networks in [89, 27], where the asymptotic SIR is shown to be approached with impractically large number of antennas, e.g. 10^4 antennas. Related work in [82] applied stochastic geometry to study the uplink interference in a massive MIMO network. A linear scaling between the numbers of base station antennas and scheduled users was found to maintain the mean interference, which need not preserve the SIR distribution.

3.2 Contributions

In this chapter, I propose a stochastic geometry framework to derive the uplink SINR and rate distributions in a large-scale cellular network using multi-user MIMO. To model the uplink topology, I propose an exclusion ball

model based on prior work [130], which simplifies the computation. Channel estimation error due to pilot contamination is also considered in the system model. The proposed framework also incorporates the fractional power control by compensating for a fraction of the path loss as in long term evolution (LTE) systems [146]. Based on the framework, I derive analytical expressions for the uplink SINR distribution for both MRC and ZF receivers in the massive MIMO regime. Unlike prior work analyzing asymptotic performance with infinity antennas [89, 27], the SINR coverage is examined as a function of the number of base station antennas and scheduled users per cell.

I apply the SINR results to investigate the interference-limited case, as numerical results show that the impact of noise becomes minor in the rban macro-cell scenario with certain typical system parameters. I derive scaling laws between the number of base station antennas and scheduled users per cell to maintain the same uplink SIR distributions. Unlike the linear scaling law examined in prior work [82, 64], I find that a super-linear scaling is generally required for MRC receivers to maintain the uplink SIR distributions, due to the near-far effect from intra-cell interference. For ZF receivers, I show that a linear scaling law still holds, as the intra-cell interference is negligible. I use the scaling law results to quantify the performance gap between ZF and MRC receivers, in terms of the difference in the number of antennas to provide the same SIR distribution. The results show that ZF receivers provides better SIR coverage than MRC receivers; the performance gap increases with the number of scheduled users in a cell, and is reduced with the fractional power control, as

it mitigates the near-far effect from intra-cell interference. Simulations verify our analysis, and indicate that the scaling laws derived from the stochastic geometry framework also apply to the hexagonal model. Numerical results on rate also show that the average per user rate can be maximized by adjusting the compensation fraction, where the optimal fraction is around 0.5 for MRC receivers, and 0.2 for ZF receivers.

Our prior work in [34] focused on the performance of MRC receivers, and provided an expression for the SIR distribution assuming no power control. In this chapter, I incorporate thermal noise in the analysis, extend the results to the case of general fractional power control for MRC receivers, and analyze the performance of ZF receivers.

This chapter is organized as follows. I present the system model for network topology and channel assumptions in Section 3.3. I analyze the performance of MRC receivers in Section 3.4, and that of ZF receivers in Section 3.5.1, followed by a performance comparison of two receivers in Section 3.5.2. I present numerical results to verify the analysis in Section 3.6, and conclude the chapter in Section 3.7.

3.3 System Model

In this section, I introduce the system model for an uplink massive MIMO cellular network. I focus on the networks operated in the sub-6 GHz band; the proposed model can be extended for massive MIMO at mmWave frequencies by incorporating key differences in propagation and hardware con-

straints [33]. Each base station is assumed to have M antennas. In each time-frequency resource block, a base station can simultaneously schedule K users in its cell. Let X_ℓ be the location of the ℓ-th base station, $Y_\ell^{(k)}$ be the location of the k-th scheduled user in the cell of ℓ-th base station, and $\mathbf{h}_{\ell\ell'}^{(k)}$ the channel vector from X_ℓ to $Y_{\ell'}^{(k)}$.

I consider a cellular network with perfect synchronization, and assume the following pilot-aided channel estimation in the uplink. In the uplink channel training stage, the scheduled users $Y_\ell^{(k)}$ send their assigned pilots $\mathbf{t}_k$, and base stations X_ℓ estimate the channels by correlating the corresponding pilots and using an MMSE estimator; in the uplink data transmission, the base stations will apply either MRC or ZF receivers, based on the channel estimates derived from uplink pilots. Further, I assume the pilots $\{\mathbf{t}_k\}_{1\leq k\leq K}$ are orthogonal and fully reused in the network. Note that the system model assumption applies to general uplink multi-user MIMO networks with pilot-aided channel estimation in the uplink, including but not limited to the TDD massive MIMO [90].

Now, I introduce the channel model assumptions. The channel is assumed to be constant during one resource block and fades independently from block to block. Moreover, I apply a narrowband channel model, as frequency selectivity in fading can be minimized by techniques like orthogonal frequency-division multiplexing (OFDM) and frequency domain equalization [59]. I express the channel vector $\mathbf{h}_{\ell n}^{(k)} \in \mathcal{C}^{M\times 1}$ as

$$\mathbf{h}_{\ell n}^{(k)} = \left(\beta_{\ell n}^{(k)}\right)^{1/2} \mathbf{\Phi}_{\ell n}^{(k)1/2} \mathbf{w}_{\ell n}^{(k)}, \tag{3.1}$$

where $\beta_{\ell n}^{(k)}$ is the large-scale path loss, $\mathbf{w}_{\ell n}^{(k)} \in \mathcal{C}^{M\times 1}$ is a Gaussian vector with the distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ for Rayleigh fading, and $\mathbf{\Phi}_{\ell n}^{(k)} \in \mathcal{C}^{M\times M}$ is the covariance matrix to account for potential correlations in small-scale fading. In this chapter, I focus on the case of IID channel fading channels, i.e., $\Phi_{\ell n}^{(k)} = \mathbf{I}_M$. The incorporation of spatial correlations in fading is deferred to future work.

The large-scale path loss gain $\beta_{\ell n}^{(k)}$ is computed as

$$\beta_{\ell n}^{(k)} = C \left(R_{\ell n}^{(k)} \right)^{-\alpha}, \tag{3.2}$$

where C is a constant determined by the carrier frequency and reference distance, $\alpha > 2$ is the path loss exponent.

Next, I introduce the network topology assumptions based on stochastic geometry. I assume the base stations are distributed as a PPP with a density λ_{b}. A user is assumed to be associated with the base station that provides the minimum path loss signal. In this chapter, each base station is assumed to serve K scheduled users that are independently and uniformly distributed in its Voronoi cell [40]. The assumption is equivalent to that in prior work [130], where the scheduled user process is obtained by (i) generating an overall user process as a PPP, and (ii) randomly selecting K out of associated users in each cell as its scheduled users, under the full buffer assumption that the overall user process is sufficiently dense, such that each base station has at least K candidate users in the cell. Without loss of generality, a typical *scheduled user* $Y_0^{(1)}$ is fixed at the origin, and its serving base station X_0 is denoted as the *tagged base station* in this chapter. I will investigate the SINR and rate

performance at this typical user.

Now I focus on modeling the distribution of *scheduled user process* in a resource block. For $1 \leq k \leq K$, the k-th scheduled user $Y_\ell^{(k)}$ in each cell is assigned with the same pilot $\mathbf{t}_k$. Let $\mathcal{N}_{\mathrm{u}}^{(k)}$ be the point process formed by the locations of the k-th scheduled users $Y_\ell^{(k)}$ from each cell. Note that the scheduled user process $\mathcal{N}_{\mathrm{u}}^{(k)}$ is non-stationary (also non-PPP), as their locations are correlated with the base station process, and the presence of one scheduled user using $\mathbf{t}_k$ prohibits the others' in the same cell [108, 130, 53]. Unfortunately, the correlations in the scheduled users' locations make the exact analysis intractable. In [130], the authors proposed an uplink model to account for the pairwise correlations, where the other-cell scheduled users for base station X_0 in $\mathcal{N}_{\mathrm{u}}^{(k)}$ is modelled as an inhomogeneous PPP with a density function of

$$\lambda_{\mathrm{u}}(r) = \lambda_{\mathrm{b}} \left(1 - \mathrm{e}^{-\lambda_{\mathrm{b}} r^2}\right), \tag{3.3}$$

where r is the distance to base station X_0. To further simplify the analysis, e.g., the computation in (3.23) and (3.26), I propose an *exclusion ball approximation*, as a first-order approximation of the model in [130], on the distribution of the scheduled user process $\mathcal{N}_{\mathrm{u}}^{(k)}$ as follows.

Assumption 3.3.1. *The following assumptions are made to approximate the exact scheduled users' process* $\mathcal{N}_{\mathrm{u}}^{(k)}$.

1. *The distances* $R_{\ell\ell}^{(k)}$ *from a user to their associated base stations are assumed to be IID Rayleigh random variables with mean* $0.5\sqrt{1/\lambda_{\mathrm{b}}}$ *[108].*

2. *The other-cell scheduled user process* $\mathcal{N}_{\mathrm{u}}^{(k)}$ *is modeled by a homogenous PPP of density* λ_{b} *outside an exclusion ball centered at the tagged base station* X_0 *with a radius* R_{e}.

3. *The scheduled users processes using different pilots* $\mathcal{N}_{\mathrm{u}}^{(k)}$ *and* $\mathcal{N}_{\mathrm{u}}^{(k')}$ *are assumed to be independent for* $k \neq k'$.

Note that in the exclusion ball model, I equivalently use a step function $\lambda_{\mathrm{b}}\left(1-\mathbb{I}(r<R_{\mathrm{e}})\right)$ to approximate the density function in (3.3), where $\mathbb{I}(\cdot)$ is the indicator function.

In this chapter, I let $R_{\mathrm{e}}=\sqrt{1/\left(\pi\lambda_{\mathrm{b}}\right)}$ by matching the average number of the excluded points from a homogenous PPP of density λ_{b} in the step function and in (3.3), i.e., by letting $\lambda_{\mathrm{b}}\pi R_{\mathrm{e}}^2=2\pi\lambda_{\mathrm{b}}\int_0^{\infty}\mathrm{e}^{-\lambda_{\mathrm{b}}\pi r^2}r\mathrm{d}r=1$. An alternative explanation for our choice of R_{e} is to let the size of the exclusion ball πR_{e}^2 equal the average cell size $1/\lambda_{\mathrm{b}}$ [24]. In Section 3.6, I show that the SINR distributions derived based on the exclusion ball assumption, as well as the approximations made in our subsequent derivation, match well with the simulation using the exact user distribution.

Fractional power control, as used in the LTE systems [146], is assumed in both the uplink training and uplink data stages: the user $Y_{\ell}^{(k)}$ transmits with power

$$P_{\ell}^{(k)}=P_{\mathrm{t}}\left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon}, \tag{3.4}$$

where $\beta_{\ell\ell}^{(k)}$ is the path loss in the corresponding signal link, $\epsilon\in[0,1]$ is the fraction of the path loss compensation, and P_{t} is the open loop transmit power

with no power control. I omit the constraint on the maximum uplink transmit power for simplicity; the constraint can be incorporated by applying the truncated channel inversion power control model [53] to determine the transmit power. I note that ignoring the maximum transmit power constraint increases the average transmit power, and reduces the impact of noise. The incorporation of more complicated power control algorithms is deferred to future work. The noise power is denoted as σ^2.

In the uplink training stage, after correlating the received training signal with the corresponding pilot, base station X_0 has an observation of the channel $\mathbf{h}_{00}^{(1)}$ as

$$\mathbf{u}_{00}^{(1)} = \sqrt{P_0^{(1)}}\mathbf{h}_{00}^{(1)} + \sum_{\ell>0} \sqrt{P_\ell^{(1)}}\mathbf{h}_{\ell 0}^{(1)} + \mathbf{n}_{\mathrm{t}},$$

where $\mathbf{n}_{\mathrm{t}}$ is the noise vector in the training stage following the distribution $\mathcal{CN}\left(\mathbf{0}, \frac{\sigma^2}{K}\mathbf{I}_M\right)$.

I assume for $\ell > 0$, the large-scale path losses $\beta_{0\ell}^{(1)}$ are perfectly known to base station X_0. Since the channels are assumed to be IID Rayleigh fading, the channel $\mathbf{h}_{00}^{(1)}$ is estimated by an MMSE estimator as

$$\bar{\mathbf{h}}_{00}^{(1)} = \frac{\sqrt{P_0^{(1)}}\beta_{00}^{(1)}}{\sum_\ell P_\ell^{(1)}\beta_{0\ell}^{(1)} + \frac{\sigma^2}{K}}\mathbf{u}_{00}^{(1)}, \tag{3.5}$$

where $\bar{\mathbf{h}}_{00}^{(1)}$ is the estimation of $\mathbf{h}_{00}^{(1)}$. Due to the orthogonality principle, the channel vector $\mathbf{h}_{00}^{(1)}$ can be decomposed as

$$\mathbf{h}_{00}^{(1)} = \bar{\mathbf{h}}_{00}^{(1)} + \hat{\mathbf{h}}_{00}^{(1)}, \tag{3.6}$$

where $\hat{\mathbf{h}}_{00}^{(1)}$ is the estimation error following $\mathcal{CN}\left(\mathbf{0}, \beta_{00}^{(1)}\left(1-\frac{P_0^{(1)}\beta_{00}^{(1)}}{\sum_\ell P_\ell^{(1)}\beta_{0\ell}^{(1)}+\frac{\sigma^2}{K}}\right)\mathbf{I}\right)$.

Let $s_\ell^{(k)}$ be the uplink data symbol for user $Y_\ell^{(k)}$ with $\mathbb{E}\left[|s_\ell^{(k)}|^2\right] = P_\ell^{(k)}$. In uplink data transmission, base station X_ℓ is assumed to use the combiner vector $\mathbf{g}_{\ell\ell}^{(k)}$ to decode $s_\ell^{(k)}$ from $Y_\ell^{(k)}$, based on the channel estimate $\bar{\mathbf{h}}_{\ell\ell}^{(k)}$. Then, at base station X_0, the decoded symbol $\hat{s}_0^{(1)}$ for the typical user $X_0^{(1)}$ is

$$\hat{s}_0^{(1)} = \mathbf{g}_{00}^{(1)*}\bar{\mathbf{h}}_{00}^{(1)}s_\ell^{(k)} + \underbrace{\mathbf{g}_{00}^{(1)*}\hat{\mathbf{h}}_{00}^{(1)}s_\ell^{(k)} + \sum_{(\ell,k)\neq(0,1)}\mathbf{g}_{00}^{(1)*}\mathbf{h}_{0\ell}^{(k)}s_\ell^{(k)} + \mathbf{g}_{00}^{(1)*}\mathbf{n}_{\mathrm{u}}}_{\text{unknown at base station}}, \tag{3.7}$$

where $\mathbf{n}_{\mathrm{u}} \in \mathcal{C}^{M\times 1}$ is the thermal noise vector in the uplink data transmission. Treating the unknown terms at base station X_0 as uncorrelated additive noise, the uplink SINR for the typical user $Y_0^{(1)}$ is

$$\mathrm{SINR} = \frac{P_0^{(1)}|\mathbf{g}_{00}^{(1)*}\bar{\mathbf{h}}_{00}^{(1)}|^2}{P_0^{(1)}\mathbb{E}|\mathbf{g}_{00}^{(1)*}\hat{\mathbf{h}}_{00}^{(1)}|^2 + \sum_{(\ell,k)\neq(0,1)}P_\ell^{(k)}\mathbb{E}|\mathbf{g}_{00}^{(1)*}\mathbf{h}_{0\ell}^{(k)}|^2 + |\mathbf{g}_{00}^{(1)}|^2\sigma^2}, \tag{3.8}$$

where the expectation operator is taken over the channel estimation error and small-scale fading in the interference links. I will investigate the SINR distributions for MRC and ZF receivers in the following sections.

The proposed system model represents a simple multi-user MIMO systems in which the SINR expression can be analyzed using stochastic geometry. In the following sections, I will study the uplink SINR and rate distributions for MRC and ZF receivers, when the number of base station antennas is large.

3.4 Performance Analysis for MRC Receivers

In this section, I derive an approximate SINR distribution in an uplink multi-user MIMO network, where the approximation becomes tight in the massive MIMO regime, e.g. when $M > 64$. Then, I focus on the interference limited case, as numerical results show that the uplink SINR is dominated by the interference in certain urban macro-cell scenarios. I derive a scaling law between the number of users and antennas that maintains the uplink SIR distribution at the typical user. Finally, I present a method to compute the per-user achievable rate and cell throughput, based on the SINR distribution.

3.4.1 SIR Coverage Analysis

Now I investigate the uplink SINR coverage based on the system model. With MRC receivers, I assume that base station X_0 applies the combining vector $\mathbf{g}_{00}^{(k)}$ as a scaled version of the channel estimate $\bar{\mathbf{h}}_{00}^{(1)}$ to decode the signal from $Y_{00}^{(1)}$:

$$\mathbf{g}_{00}^{(1)} = \frac{\sum_{\ell} P_{\ell}^{(1)} \beta_{0\ell}^{(1)} + \frac{\sigma^2}{K}}{\sqrt{P_0^{(1)} \beta_{00}^{(1)}}} \bar{\mathbf{h}}_{00}^{(1)} = \mathbf{u}_{00}^{(1)}. \tag{3.9}$$

Note the scaling on the combining vector is intended to simplify expressions, and will not change the SINR distribution. Then, using the combining vector in (3.9), the SINR expression can be simplified in (3.10) as

$$\frac{(M+1)\left(\beta_{00}^{(1)}\right)^{2(1-\epsilon)}}{M\Delta_2^{(1)} + \left(\beta_{00}^{(1)}\right)^{1-\epsilon}\Delta_1^{(1)} + \left(\sum_{k=2}^{K}\left(\beta_{00}^{(k)}\right)^{1-\epsilon} + \sum_{k=1}^{K}\Delta_1^{(k)}\right)\left(\left(\beta_{00}^{(1)}\right)^{1-\epsilon} + \Delta_1^{(1)}\right)} \tag{3.10}$$

where $\Delta_1^{(k)} = \sum_{\ell>0} \left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon} \beta_{0\ell}^{(k)} + \frac{\sigma^2}{KP_{\mathrm{t}}}$, and $\Delta_2^{(k)} = \sum_{\ell>0} \left(\beta_{\ell\ell}^{(k)}\right)^{-2\epsilon} \left(\beta_{0\ell}^{(k)}\right)^2$. The derivation to obtain (3.10) is given in Section 3.8.1. Note that $\Delta_1^{(k)}$ and $\Delta_2^{(k)}$ correspond to the sum of certain interference terms from other-cell users.

Next, I denote the exact SINR distribution for (3.10) (using the exact scheduled user distribution defined in Section 3.3 but not the exclusion ball assumption) as $\mathbb{P}(\mathrm{SINR} > T)$. Due to pilot contamination, the combining vector $\mathbf{g}_{00}^{(k)}$ is correlated with certain interference channel vectors as shown in (3.9). As a result, the denominator in (3.10) contains cross-products of the path losses from different interferers. Moreover, different cross-product terms in the denominator of (3.10) can be correlated, as they may contains common path loss terms, which renders the exact derivation of $\mathbb{P}(\mathrm{SINR} > T)$ intractable. Therefore, I compute an approximate SINR distribution $\bar{\mathbb{P}}(\mathrm{SINR} > T)$, which I argue in Section 3.6 is a good match for $\mathbb{P}(\mathrm{SINR} > T)$, in Theorem 3.4.1.

Theorem 3.4.1 (MRC SINR)**.** *In the proposed massive MIMO networks, an approximate uplink SINR distribution with MRC receivers can be computed as*

$$\bar{\mathbb{P}}(\mathrm{SINR} > T) = \sum_{n=1}^{N} \binom{N}{n} (-1)^{n+1} \int_0^\infty \mathrm{e}^{-t - \ell T\eta C_1 t^{\alpha(1-\epsilon)} - \ell T\eta C_2 t^{\frac{\alpha}{2}(1-\epsilon)}} C_3(t)\mathrm{d}t, \tag{3.11}$$

where N is the number of terms used in the calculation, $\eta = N(N!)^{-\frac{1}{N}}$, $C_{\sigma^2} = \frac{\sigma^2}{KP_{\mathrm{t}}C^{1-\epsilon}(\lambda_{\mathrm{b}}\pi)^{\frac{\alpha(1-\epsilon)}{2}}}$, $C_1 = \frac{K+1}{M+1}\left(\frac{2\Gamma^\alpha(\frac{\epsilon}{2}+1)}{(\alpha-2)} + C_{\sigma^2}\right)$, $C_2 = \frac{M\Gamma^\alpha(\epsilon+1)}{(M+1)(\alpha-1)} + \frac{K}{M+1}\left(\frac{2\Gamma^\alpha(\frac{\epsilon}{2}+1)}{(\alpha-2)} + C_{\sigma^2}\right)^2$,

$$C_3(t) = \left(\int_0^\infty \mathrm{e}^{-u - u^{-\frac{\alpha}{2}(1-\epsilon)} \ell T\eta C_4(t)}\mathrm{d}u\right)^{K-1}$$

$$\approx \left(1 - \ell\eta T C_4(t) \int_0^\infty \frac{\mathrm{e}^{-u}}{\ell T \eta C_4(t) + u^{-\frac{\alpha}{2}(1-\epsilon)}} \mathrm{d}u\right)^{K-1},$$

$C_4(t) = \frac{1}{M+1}\left(\left(\frac{2\Gamma^\alpha(\frac{\epsilon}{2}+1)}{\alpha-2} + C_{\sigma^2}\right) t^{\alpha(1-\epsilon)} + t^{\frac{\alpha}{2}(1-\epsilon)}\right)$, *and* $\Gamma(\alpha) = \int_0^\infty \mathrm{e}^{-t} t^{\alpha-1} \mathrm{d}t$ *is the gamma function.*

Proof. See Section 3.8.2. □

Besides the exclusion ball approximation, the main approximation in Theorem 3.4.1 is to replace certain out-of-cell interference terms by their means in (3.23) and (3.26). The approximation results in a minor error in the SINR distribution, as (i) with K users in a cell, the intra-cell interference dominates the out-of-cell interference with high probability; (ii) with large antenna arrays, the ratio of the signal power to certain out-of-cell interference power terms, e.g. the terms in $\Delta_1^{(k)}$, decays as $\frac{1}{M}$. In Section 3.6, using $N \geq 5$ terms, the distribution $\bar{\mathbb{P}}(\mathrm{SINR} > T)$ computed in Theorem 3.4.1 is shown to be a good match with the SINR distribution $\mathbb{P}(\mathrm{SINR} > T)$ from Monte Carlo simulations. In addition, the error of the approximation becomes more prominent with a smaller noise power, as all the approximations are made with respect to the interference distribution. The expression is intended for the massive MIMO regime when $M \gg 1$, as the error of the approximations decays with $\frac{1}{M}$. In simulations, I find that the results in the theorem generally applies to the multi-user MIMO networks with not-so-large M, e.g. the case of $(M, K) = (10, 2)$.

In Theorem 3.4.1, the noise power is taken account by the parameter $C_{\sigma^2} = \frac{\sigma^2}{K P_{\mathrm{t}} C^{1-\epsilon} (\lambda_{\mathrm{b}} \pi)^{\frac{\alpha(1-\epsilon)}{2}}}$, which shows that the impact of noise on the SINR

is reduced with a larger number of scheduled users per cell K, a higher base station density λ_{b}, a smaller path loss α, and a larger power control parameter ϵ. Besides, the impact of noise goes down with larger M, as in the expressions for C_1 and C_2, the noise parameter C_{σ^2} is divided by $(M+1)$.

Next, I focus on the performance of interference-limited networks. I will show in Section 3.6 that the impact of noise is negligible in certain urban macro-cell cases with $M = 64$ antennas at base stations. Then, the general expression in Theorem 3.4.1 can be further simplified in the following special cases.

Case 1 (Full power control, $\epsilon = 1$): In this case, the transmitting power at scheduled user $Y_\ell^{(k)}$ is adjusted to compensate for the full path loss, i.e., $P_\ell^{(k)} = P_{\mathrm{t}}\beta_{\ell\ell}^{(k)}$, such that a base station receives equal signal powers from all of its associated users. When $\epsilon = 1$, the SIR distribution can be simplified as in the following corollary.

Corollary 3.4.1.1. *With $\epsilon = 1$ and $\sigma^2 = 0$, the approximate SIR distribution can be computed as*

$$\bar{\mathbb{P}}(\mathrm{SIR} > T) = \sum_{n=1}^{N} \binom{N}{n} (-1)^{n+1} \mathrm{e}^{-T\eta\ell\left(\frac{C_5 K + \Gamma^{\alpha}(1.5)}{M+1} + \frac{1}{\alpha-1}\right)}, \tag{3.12}$$

where $C_5 = \frac{4\Gamma^{2\alpha}(1.5) + (\alpha^2-4)\Gamma^{\alpha}(1.5)}{(\alpha-2)^2}$.

Based on Corollary 3.4.1.1, a linear scaling law between the number of users and antennas is observed as follows.

Corollary 3.4.1.2. *With $\epsilon = 1$ and $\sigma^2 = 0$, to maintain the uplink SIR distribution unchanged, the scaling law between the number of base station antennas M and users per cell K is approximately*

$$(M+1) \sim \left(K + \frac{\Gamma^{\alpha}(1.5)}{C_5} \right) \approx K. \tag{3.13}$$

Note that when $\epsilon = 1$, the linear scaling law matches prior results in [64, Sec. IV], where the path loss to all associated users in the typical cell was assumed to be identical. The linear scaling law, however, does not apply to other cases with $\epsilon < 1$, e.g. in the following case without power control.

Case 2 (No power control, $\epsilon = 0$): In this case, the fraction of the path loss compensation is $\epsilon = 0$. Then, the uplink SIR can be evaluated as follows.

Theorem 3.4.2. *With $\epsilon = 0$ and $\sigma^2 = 0$, an approximate uplink SIR distribution can be calculated as*

$$\bar{\mathbb{P}}(\mathrm{SIR} > T) = \sum_{n=1}^{N} \binom{N}{n} (-1)^{n+1} \int_0^{\infty} \mathrm{e}^{-(\mu\Gamma(1-2/\alpha)(n\eta T)^{2/\alpha}+1)t - \frac{n\eta T}{\alpha-1} t^{\alpha}} \mathrm{d}t, \tag{3.14}$$

where N is the number of terms used in the computation, and $\mu = \frac{K}{(M+1)^{2/\alpha}}$.

Proof. The proof is similar to that in [34, Appendix A]. □

I will show in Section 3.6 that Theorem 3.4.2 provides a tight approximation of the exact SIR distribution $\mathbb{P}(\mathrm{SIR} > T)$, when $N \geq 5$ terms are used. Moreover, note that in (3.14), the number of antennas M and the number of

scheduled users per cell K only affect the value of μ. Therefore, by Theorem 3.4.2, in the no power control case, I observe the following scaling law to maintain SIR.

Corollary 3.4.2.1. *Assuming no power control, the approximate scaling law to maintain the same uplink SIR distribution is* $(M+1) \sim K^{\alpha/2}$*, which is a superlinear polynomial scaling when* $\alpha > 2$.

In the case of no power control, the difference in the path losses between the typical user and the intra-cell interferers affect the SIR distribution, and thus the scaling law to maintain the SIR becomes a function of the path loss exponent. The super-linearity in the scaling law can be explained by the near-far effect of the intra-cell interference from multiple users in a cell. With no power control, the cell edge users will receive weaker signals than the cell center user. With a uniform user distribution in a cell, the typical user will be more likely to be located at the cell edge. When increasing the number of scheduled users K in a cell, the probability that the interference from a cell-center interferer dominates the signal from the typical user increases. Therefore, compared with the linear scaling law with full power control ($\epsilon = 1$) where such near-far effect is mitigated, more antennas will be needed in the no power control case to reduce the intra-cell interference, and preserve the SIR distribution, when increasing K.

Next, I focus on the scaling law in the general fractional power control case with $\epsilon \in (0, 1)$. It is difficult to derive the exact scaling law directly from

the expression (3.11), due to the integral form. Since with the fractional power control, the equivalent path loss exponent in the signal link linearly scales with ϵ, I propose the following approximate scaling law by linearly fitting the exponent s of the scaling law $(M+1) \sim K^s$, based on two special cases of ϵ: by Corollary 3.4.1.2, when $\epsilon = 1$, $s = 1$; and by Theorem 3.4.2, when $\epsilon = 0$, $s = \frac{\alpha}{2}$. Therefore, for general $0 < \epsilon < 1$, the linearly fitted exponent of the scaling law s is given as follows.

Scaling law 3.4.1. *With fractional power control, the scaling law between M and K is approximately $(M+1) \sim K^s$, where the exponent of the scaling law is $s = \frac{\alpha}{2}(1-\epsilon) + \epsilon$.*

Scaling law 3.4.1 reveals that a (superlinear) polynomial scaling law between K and M is required to maintain uplink SIR distribution, for a general $\epsilon < 1$. The results in Scaling law 3.4.1 are verified by numerical simulations in Section 3.6.

3.4.2 Rate Analysis

In this section, I apply the SINR results to compute the achievable rate. First, I define the average achievable spectrum efficiency at a typical user as

$$\tau_0 = \mathbb{E}\left[\log_2\left(1 + \min\{\text{SINR}, T_{\max}\}\right)\right], \tag{3.15}$$

where $T_{\max}$ is a SINR distortion threshold determined by limiting factors like distortion in the radio frequency front-end. By [29, Section III-C], given the SINR distribution $\mathbb{P}(\text{SINR} > T)$, the average achievable spectrum efficiency

can be computed as $\tau_0 = \frac{1}{\ln(2)} \int_0^{T_{\max}} \frac{\mathbb{P}(\text{SINR}>x)}{1+x} \mathrm{d}x$. To take account for the overhead, let ψ be the fraction time for overhead. In this chapter, for simplicity, I only consider the overhead due to uplink channel training, and compute the overhead fraction ψ as $\psi = \frac{T_\mathrm{t}}{T_\mathrm{c}} = \frac{K}{T_c}$, where T_t and T_c are the length of channel training period and coherent time, in terms of the number of symbol time. The length of channel training is assumed to be equal to the number of scheduled users in a cell, as I assumed full reuse of orthogonal pilots throughout the network. Then the average achievable rate with the overhead penalty $\bar{\tau}_0$ equals

$$\bar{\tau}_0 = K(1-\psi)\tau_0, \tag{3.16}$$

Note that when ignoring thermal noise, the scaling law to maintain SINR distribution also maintains the average achievable rate τ_0. When taking account for the training overhead penalty ψ, however, the scaling law will not keep $\bar{\tau}_0$ unchanged, as $1-\psi$ linearly decreases with K, unless ψ is negligible, e.g. when the coherence time $T_\mathrm{c} \gg K$. Next, I define the average cell throughput τ_{cell}, in terms of spectrum efficiency, as

$$\tau_{\text{cell}} = K(1-\psi)\tau_0. \tag{3.17}$$

I will examine the average cell throughput as a function of M and K in Section 3.6. Before that, I continue to present the results for ZF receivers in the next section.

3.5 Performance Analysis with ZF Receivers

In this section, I will investigate the performance of ZF receivers in IID fading channels. First, I derive the SINR and rate distributions with ZF receivers. Then, I apply the analytical results to compare the performance of MRC and ZF receivers in an interference-limited network. In particular, I aim to answer the question: compared with MRC receivers, how many antennas can be saved by applying ZF receivers, while keeping the same uplink SIR distribution.

3.5.1 SINR Analysis of ZF Receivers

Now I begin to investigate the performance of ZF receivers in an uplink massive MIMO network. For ZF receivers, I still focus on the case of IID fading. To cancel the intra-cell interference, base station X_ℓ will apply the combining vector $\mathbf{g}_{\ell\ell}^{(k)}$ for user $Y_\ell^{(k)}$ as

$$\mathbf{g}_{\ell\ell}^{(k)} = \mathbf{H}_\ell^\dagger[:, k], \tag{3.18}$$

where $\mathbf{H}_\ell = \left[\mathbf{u}_{\ell\ell}^{(1)}, \mathbf{u}_{\ell\ell}^{(2)}, \ldots, \mathbf{u}_{\ell\ell}^{(K)}\right] \in \mathcal{C}^{M\times K}$ is the matrix of all estimated channels to the associated users in cell X_ℓ, and $\mathbf{u}_{\ell\ell}^{(k)} = \frac{\sum_{\ell'} P_{\ell'}^{(1)} \beta_{\ell\ell'}^{(1)} + \frac{\sigma^2}{K}}{\sqrt{P_\ell^{(k)} \beta_{\ell\ell}^{(k)}}} \bar{\mathbf{h}}_{\ell\ell}^{(k)}$ is a scaled version of the channel estimate. The scaling in the channel estimates will not change the uplink SINR distribution, as it will only cause certain scaling in the corresponding combining vector. Similar to the case of MRC receivers, the exact uplink SINR distribution is difficult to derive, as due to pilot contamination, the combining vector is correlated with certain interference channel

vectors. Therefore, applying the same approximations in (3.23) and (3.26), I derive an approximate distribution for the uplink SINR expression in (3.8) for the typical user $X_0^{(1)}$ in the following theorem.

Theorem 3.5.1. *With $M \gg K$ and ZF receivers, an approximate uplink SINR distribution for the typical user can be calculated by*

$$\bar{\mathbb{P}}(\mathrm{SINR} > T) = \sum_{n=1}^{N} \binom{N}{n} (-1)^{n+1} \int_0^{\infty} \mathrm{e}^{-n\eta T\left(C_6 t^{\frac{\alpha}{2}(1-\epsilon)} + C_7 t^{\alpha(1-\epsilon)}\right) - t} \mathrm{d}t, \quad (3.19)$$

where the constant $C_6 = C_9 \left(\frac{1}{M-K+1} + \frac{1}{M+1} + \frac{M(K-1)}{(M-K+1)^2}\right) + \frac{M(K-1)C_8}{(M-K+1)^2}$,

$$C_7 = \frac{M}{M+1}\frac{\Gamma^{\alpha}(\epsilon+1)}{\alpha-1} + \left(\frac{1}{M+1} + \frac{(K-1)M}{(M-K+1)^2}\right) C_9^2 + \frac{(K-1)M}{(M-K+1)^2} C_8 C_9,$$

$C_8 = \frac{2\Gamma^{\alpha}(\frac{\epsilon}{2}+1) + (\alpha-2)C_{\sigma^2}}{(\alpha-2)(1+C_{\sigma^2}) + 2\Gamma^{\alpha}(\frac{\epsilon}{2}+1)}$, $C_9 = \frac{2\Gamma^{\alpha}(\frac{\epsilon}{2}+1)}{\alpha-2} + C_{\sigma^2}$, $C_{\sigma^2} = \frac{\sigma^2}{KP_{\mathrm{t}}C^{1-\epsilon}(\lambda_{\mathrm{b}}\pi)^{\frac{\alpha(1-\epsilon)}{2}}}$, N *is the number of terms used in the computation, and* $\eta = N(N!)^{-\frac{1}{N}}$.

Proof. See Section 3.8.3. □

Note that when $K = 1$, the SINR distribution in (3.19) for ZF receivers is the same as that for MRC receivers in (3.12). I will verify the tightness of the approximation $\mathbb{P}(\mathrm{SINR} > T) \approx \bar{\mathbb{P}}(\mathrm{SINR} > T)$ by numerical simulation in Section 3.6. I have the following remark on the applicable regime for Theorem 3.5.1.

Remark 3.5.1. I need the condition $M \gg K$ in the proof, as the error in the approximation in (3.27) decays as $\frac{1}{M-K+1}$. In numerical simulations, I find that the approximate SINR distribution in Theorem 3.5.1 shows a good

match with the simulations when $\frac{M}{K} \geq 3$ with $M \geq 10$. The same comment applies to Scaling law 3.5.1 below.

Next, I focus on the interference-limited case. Based on Theorem 3.5.1, I can derive an approximate scaling law between M and K to maintain the SIR distribution in the region of $M \gg K$ as follows.

Scaling law 3.5.1. *With ZF receivers and $\sigma^2 = 0$, the uplink SIR distribution of the typical user remains approximately unchanged when the number of antennas M linearly scales with the number of users per cell K as $(M+1) \sim K$.*

Proof. Note that when $\sigma^2 = 0$, $C_\sigma = 0$. The dependence on M and K in (3.19) only occurs in the constants C_6 and C_7. Therefore, it is sufficient to show that a linear scaling between M and K (approximately) maintains the values of C_6 and C_7. Note that when $M \to \infty$, and $M \gg K$, the following limits hold: $\frac{1}{M+1} \to 0$, $\frac{1}{M-K+1} \to 0$, and $\frac{M}{M+1} \to 1$. Therefore, it follows that when keeping $\frac{K}{M+1} = t$, $\lim_{M\to\infty} C_6 = \frac{(C_8+C_9)t}{1-t}$, and $\lim_{M\to\infty} C_7 = \frac{\Gamma^\alpha(\epsilon+1)}{\alpha-1} + \frac{t}{1-t}\left(\frac{4\Gamma^{2\alpha}(\frac{\epsilon}{2}+1)}{(\alpha-2)^2} + \frac{2C_8\Gamma^\alpha(\frac{\epsilon}{2}+1)}{\alpha-2}\right)$, which are invariant when $(M+1)$ linearly scales with K. □

Compared with MRC receivers, the near-far effect for users in a cell becomes minor with ZF receivers, as the intra-cell interference is largely suppressed. Therefore, a linear scaling law applies for ZF receivers even without power control. Based on the SINR coverage results, the achievable rate per user and sum throughput can be computed following the same line as in Section 3.4.2. In the next section, I will use the derived results to compare the

SIR coverage performance between MRC and ZF receivers in an interference-limited network.

3.5.2 Comparison of SIR Coverage Performance

Now assuming the network is interference-limited, I compare the SIR coverage between ZF and MRC receivers. Prior work [64] showed that ZF and MRC receivers have the same asymptotic performance, both which are limited by the pilot contamination. The analysis in [82] showed that by suppressing intra-cell interference, which turns to be more dominant than the out-of-cell interference, the ZF receivers suffers from less interference than MRC receivers. In this section, I make a quantitative comparison by answering the following question: in IID fading channels, how many base station antennas M_{ZF} is needed for ZF receivers to provide the same uplink SIR distribution as MRC receivers with M_{MRC} antennas?

Based on Scaling law 3.4.1 and Scaling law 3.5.1, I have the following proposition to determine M_{ZF} to match the SIR coverage with MRC receivers.

Proposition 3.5.1. *Assuming $M_{\mathrm{ZF}} \gg K$, ZF receivers with $(M_{\mathrm{ZF}}+1) = \xi(M_{\mathrm{MRC}}+1)$ antennas approximately provide the same uplink SIR distribution as MRC receivers with M_{MRC} antennas in a massive MIMO networks, where the scaling factor $\xi = K^{-(\frac{\alpha}{2}-1)(1-\epsilon)}$, and K is the number of scheduled users in a cell.*

Proof. For the ease of notation, let $\mathrm{ZF}(M,K)$ and $\mathrm{MRC}(M,K)$ denote the

uplink SIR distributions with ZF and MRC receivers of M antennas, when serving K users in a cell. By Scaling law 3.5.1, when $M_{\mathrm{ZF}} \gg K$, $\mathrm{ZF}(M_{\mathrm{ZF}}, K) \approx \mathrm{ZF}(\frac{M_{\mathrm{ZF}}+1}{K}, 1)$. Next, note that when $K = 1$, i.e., with a single scheduled user in a cell, MRC and ZF receivers provide the same SIR coverage. Thus, it follows that $\mathrm{ZF}(M_{\mathrm{ZF}}, K) = \mathrm{ZF}(\frac{M_{\mathrm{ZF}}+1}{K}, 1) = \mathrm{MRC}(\frac{M_{\mathrm{ZF}}+1}{K}, 1)$. Last, by Scaling law 3.4.1, $\mathrm{ZF}(M_{\mathrm{ZF}}, K) \approx \mathrm{MRC}(\frac{M_{\mathrm{ZF}}+1}{K}, 1) \approx \mathrm{MRC}((M_{\mathrm{ZF}}+1)K^{(\frac{\alpha}{2}-1)(1-\epsilon)} - 1, K)$. □

The condition $M_{\mathrm{ZF}} \gg K$ in the proposition is required to ensure the applicability of Scaling law 3.5.1. In numerical simulations, the result is found to be a good approximation with $\frac{M_{\mathrm{ZF}}}{K} > 3$. Note that the exponent of the scaling factor $-(\frac{\alpha}{2}-1)(1-\epsilon)$ is non-positive, which indicates I need $M_{\mathrm{MRC}} \geq M_{\mathrm{ZF}}$ to provide the same SIR coverage. Further, the scaling factor ξ increases with the number of the scheduled user K, which reveals that the performance gap between MRC and ZF receivers grows with K. When K increases, the mitigation of the intra-cell interference from $(K-1)$ users by ZF receivers becomes more prominent to improve SIR coverage. In addition, Proposition 3.5.1 also shows that in terms of the SIR distribution, the performance gap reduces with larger ϵ in the power control scheme, as the scaling factor ξ is a decreasing function of ϵ. Simulations show that with $\epsilon = 1$, only a minor gap exists between the SIR coverage curves for ZF and MRC receivers.

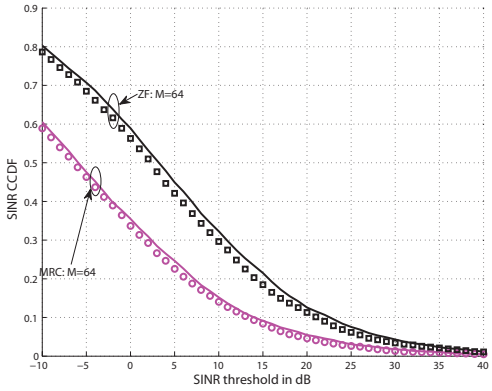
Last, I note that Proposition 3.5.1, which is drawn based on the SIR distribution, need not extend to a general SINR distribution that is not dominated by interference; prior work [147] showed that when the noise is not

negligible, MRC receivers would provide a comparable or even better SINR, compared with the ZF receivers. In the following section, I will present numerical results to validate our analytical results.

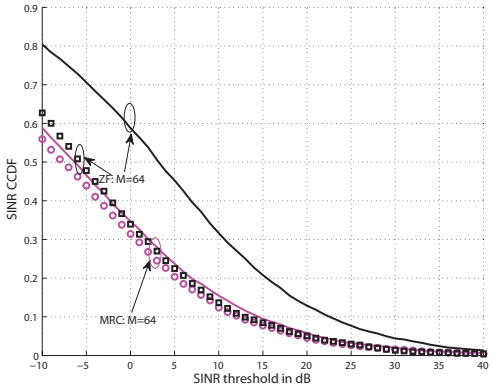
3.6 Numerical Results

In this section, I verify our analytical results with numerical simulations, which follow the procedure as: (1) generating the base station process as a PPP of density λ_{b}; (2) generating the overall user process as a PPP of density $\lambda_{\mathrm{u,o}}$, where I use $\lambda_{\mathrm{u,o}} = 60\lambda_{\mathrm{b}}$, unless otherwise specified; (3) associating the points in the overall user process to base stations, based on the minimal path loss rule, and then randomly scheduling K out of the associated users in each cell as their scheduled users; (4) picking the base station closest to the origin as the tagged base station X_0, and its first scheduled user $Y_0^{(1)}$ as the typical user; (5) generating channel vectors as IID Gaussian vectors, and computing the SINR for the iteration; (6) repeating the step (1)-(5) for 10,000 iterations, and computing the empirical distribution of the SINR at $Y_0^{(1)}$. For the simulations using hexagonal grids, I follow the same procedure except that the base station process is generated as a 19-cell hexagonal grid, and the tagged base station is the center cell. In addition, I will use $N = 5$ terms when evaluating the analytical expressions.

Impact of the noise: To begin with, I examine the impact of noise by comparing the SINR and SIR distributions for both MRC and ZF receivers in different scenarios in Fig. 3.1. In the simulations, I assume P_{t}=23 dBm,



(a) ISD=500 meters.



(b) ISD=1000 meters.

Figure 3.1: Comparison of SINR and SIR distributions. In the figures, I use markers to represent SINR curves, solid lines for SIR. I assume $K = 10$ users per cell, $\epsilon = 0$, and $\alpha = 4$ in all cases. The gap between the SIR and SINR distributions becomes minor when ISD=500 meters, which is the typical size for the urban macro cells [1].

and the bandwidth is 20 MHz as in the current LTE standards [1]. I examine the case of $\epsilon = 0$, which maximizes the impact of the noise. I simulate with two inter-site distances (ISDs): an average ISD of 500 meters in Fig. 3.1(a),

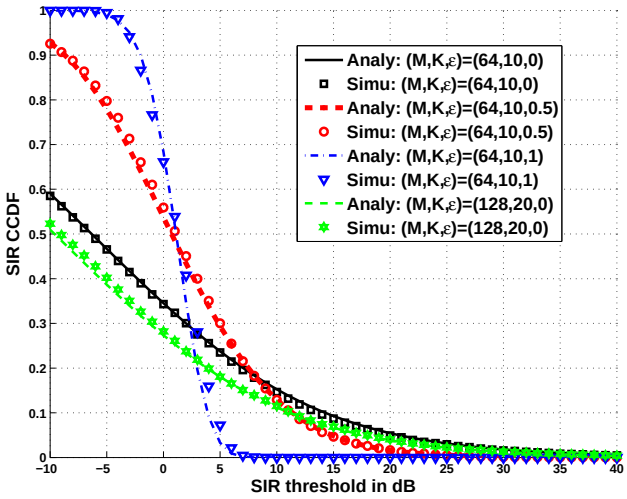


Figure 3.2: SIR coverage for MRC receivers. In the simulations, I assume $\alpha = 4$. The analytical curves are drawn based on Theorem 3.4.1, which are shown a good fit with simulation. The difference in the curves for $(M, K, \epsilon) = (64, 10, 0)$ and $(M, K, \epsilon) = (128, 20, 0)$ indicates that linear scaling between M and K does not generally preserve SIR for MRC receivers.

and 1000 meters in Fig. 3.1(b). Note that a typical ISD of 500 meters is assumed for urban macro-cells in the 3GPP standards [1]. In Fig. 3.1(a), the network with ISD=500 meters is shown to be interference-limited with $M = 64$ antennas for both MRC and ZF receivers, as the SIR curves almost coincide with the SINR curves, which justifies the interference-limited assumption in urban marco cells. In the sparse network with ISD=1000 meters, however, simulations show that even with M=64 antennas, notable gaps exist between the SINR and SIR distributions, especially for ZF receivers. In addition, the results in Fig. 3.1(b) shows that when the noise power is high, even with no power control, ZF and MRC receivers have a similar SINR coverage performance, which indicates that the SIR comparison results in Proposition 3.5.1

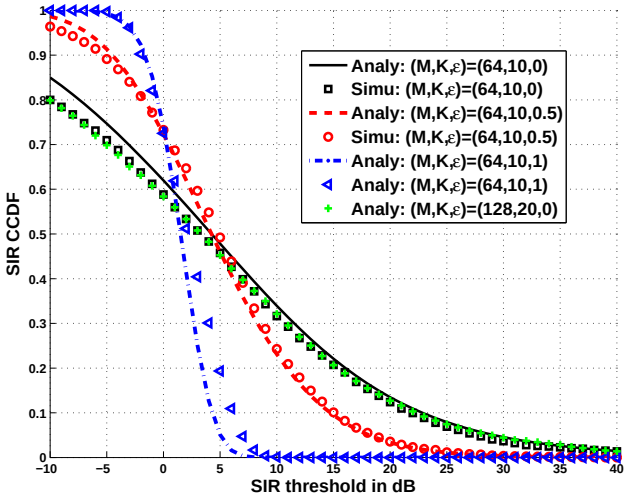


Figure 3.3: SIR distributions with ZF receivers. I assume $\alpha = 4$, and IID fading channel. The analytical curves are plotted based on Theorem 3.5.1. Simulations verify the analytical results, and show that when both M and K double, the SIR curves remain almost unchanged.

need not extends to general SINR comparisons.

SIR coverage for MRC receivers: In Fig. 3.2, I verify the analytical results for the SIR distribution with MRC receivers. Numerical results show that the SIR coverage is sensitive to the compensation fraction ϵ in the fractional power control: a large compensation fraction ϵ improves the SIR coverage in the low SIR regime at the expense of sacrificing the coverage in the high SIR regime. Besides, a comparison of the curves for $(M, K) = (64, 10)$ and $(M, K) = (128, 20)$ shows that the linear scaling law does not maintain the SIR distribution when $\epsilon = 0$.

SIR coverage for ZF receivers: I verify the analysis for ZF receivers in Fig. 3.3. The analytical curves generally match well with numer-

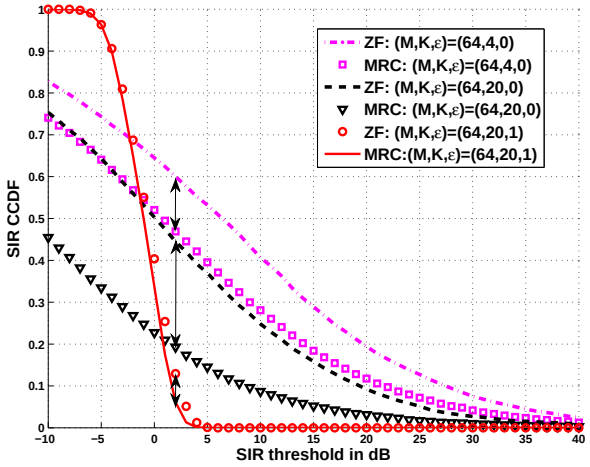


Figure 3.4: Comparison of SIR coverage with MRC and ZF receivers. I assume $\alpha = 4$. As the double arrays display, when fixing $\epsilon = 0$, the performance gap in SIR coverage is shown to increase with K; when fixing $K = 20$, the gap diminishes when $\epsilon \to 1$.

ical simulations. A comparison of the curves for $(M, K, \epsilon) = (64, 10, 0)$ and $(M, K, \epsilon) = (128, 20, 0)$ shows that unlike the case of MRC receivers, a linear scaling law between M and K maintains the SIR distribution, even when there is no fractional power control implemented.

SIR comparison between MRC and ZF receivers: I compare the uplink SIR distributions for MRC and ZF receivers in Fig. 3.4. Simulations show that ZF receivers provide better SIR coverage, due to the suppression of intra-cell interference. Moreover, for the same ϵ, the performance gap between MRC and ZF receivers increases with the number of scheduled users K, as the strength of total intra-cell interference also increases with K. When fixing M and K, the performance gap decreases with ϵ; when $\epsilon = 1$, the SIR coverage

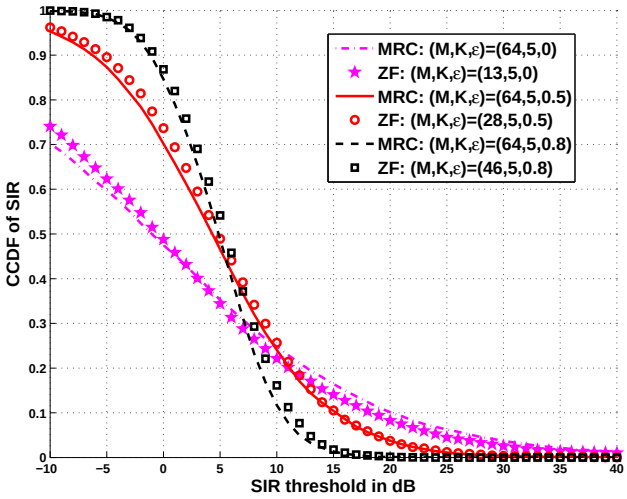
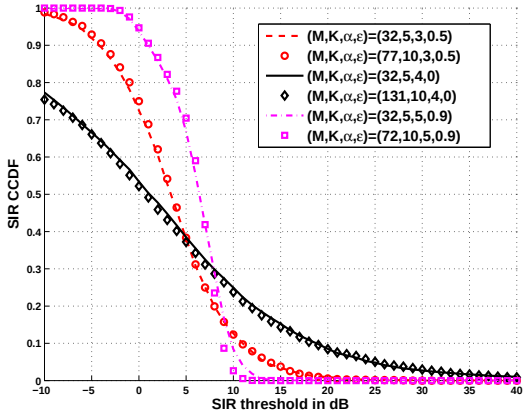


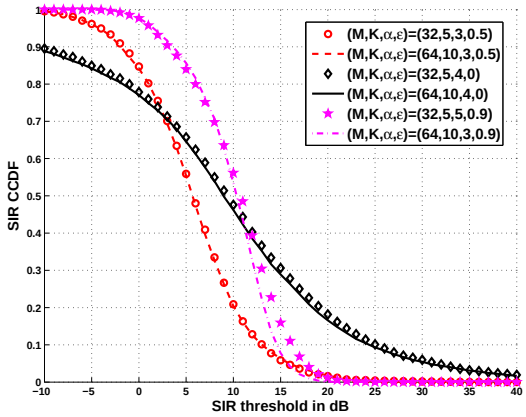
Figure 3.5: Verification of Proposition 3.5.1. In the simulation, $\alpha = 4$. In the simulation, I use the SIR curve of MRC$(64, 5)$ as a baseline for comparison. I use Proposition 3.5.1 to compute the required number of antennas for ZF receivers, to have the SIR distribution of the baseline curve.

gap becomes minimal between MRC and ZF receivers. With full compensation of path loss in power control, linear scaling laws between M and K apply to both MRC and ZF receivers, as the near-far effect for users in a cell is mitigated. When $M \gg K$, the difference in the average (residue) intra-cell interference between MRC and ZF receivers becomes minor, as it decays with $\frac{1}{M}$.

In Fig. 3.5, I verify our theoretical results in Proposition 3.5.1. In the simulation, I fix the number of antennas for the MRC receivers to be $M_{\text{MRC}} = 64$, and use Proposition 3.5.1 to calculate the required M_{ZF}, to maintain the same SIR distribution. Numerical results show a good match with our analysis; the minor mismatch in the case $\epsilon = 0$ is because Proposition



(a) SIR for MRC receivers in the hexagonal grid model.



(b) SIR for ZF receivers in the hexagonal grid model.

Figure 3.6: Verification of the scaling laws in the hexagonal model. I use $(M, K) = (32, 5)$ as the baseline curves. When increasing the number of users to K=10, I compute the required M to preserve the SIR distribution as baseline curves, according to Scaling law 3.4.1 and Scaling law 3.5.1. Simulations indicates that the scaling law results apply to the hexagonal model.

3.5.1 theoretically requires $\frac{M_{\text{ZF}}}{K} \gg 1$, while I use $\frac{M_{\text{ZF}}}{K} = \frac{13}{5}$ in the simulation.

Verification with hexagonal grid model: I verify the scaling laws

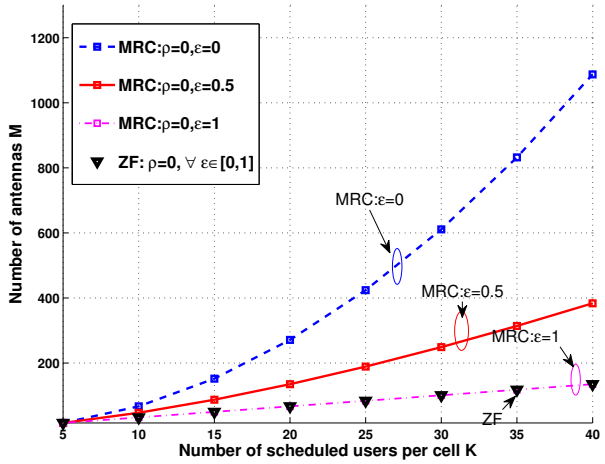


Figure 3.7: Comparison of different scaling laws. I plot the required number of antennas to provide the same SIR as that of the case $(M, K) = (16, 5)$ as a function of K with different system parameters.

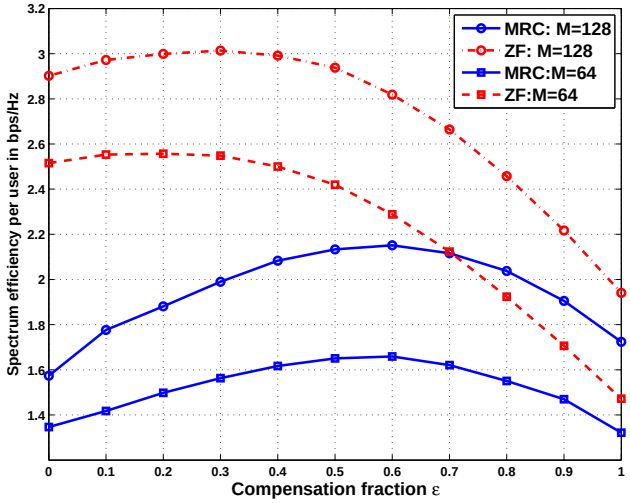
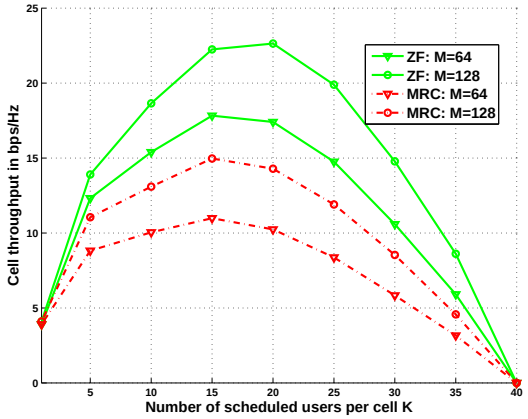
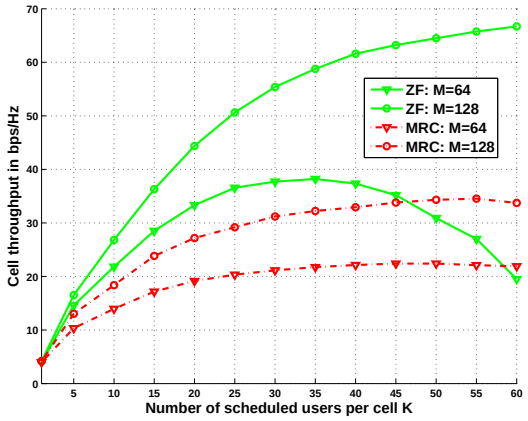


Figure 3.8: Average spectrum efficiency per user in an interference-limited network. In the simulation, I assume $T_{\max} = 21$ dB, which sets the maximum spectrum efficiency per data stream as 7 bps/Hz. Training overhead is not taken account in this figure.



(a) Cell throughput when T_c=40 symbols.



(b) Cell throughput when T_c=200 symbols.

Figure 3.9: Uplink cell throughput as a function of K. The overhead due to channel training is taken account in the simulations. I simulate an interference-limited network with ISD=500 meters. I use $\epsilon = 0.5$ for MRC receivers, and $\epsilon = 0.2$ for ZF receivers, which are shown to optimize the per user rate.

derived from stochastic geometry with the hexagonal grid model in Fig. 3.6. In the simulations, I use a layout of 19 hexagonal cells with inter-site distance of 300 meters; only the scheduled users in the central cell are counted for the SIR statistics, to avoid edge effect. In Fig. 3.6(a), for MRC receivers, I use

Table 3.1: Coherence time T_c in the examples

Mobility	Max. velocity	Max. Doppler f_{D}	Coherence time T_{c}
High	50 Km/h	92.6 Hz	40 Symbols
Low	10 Km/h	18.5 Hz	200 Symbols

a $(M, K) = (32, 5)$ as the baseline curve for comparison. When doubling the number of scheduled users to $K = 10$, I use Scaling law 3.4.1 to compute the required M to maintain the same SIR distribution, which is shown to be almost accurate with extensive combinations of the system parameters in the hexagonal grid model. Similarly, results in Fig. 3.6(b) verifies the linear scaling law for ZF receivers in Scaling law 3.5.1. This indicates that the stochastic geometry model provides reasonable predictions even for the hexagonal model.

Comparison of scaling laws: I compare scaling laws to maintain the uplink SIR distribution in different scenarios in Fig. 3.7. I plot the required number of antennas to maintain the same SIR distribution as that in the case of $(M, K) = (16, 5)$, as a function of K. As shown in the plot, for MRC receivers, given the path loss exponent α, the slope of the scaling law is determined by the fraction of path loss compensation ϵ: the linear scaling law proposed in prior work [82, 64] is only achieved when $\epsilon = 1$. Although the choice of $\epsilon = 1$ makes the system with MRC receivers linearly scalable, it need not maximize the per-user rate, as will be shown in Fig. 3.8. On the contrary, for ZF receivers, the linear scaling applies for all $\epsilon \in [0, 1]$.

Rate performance: I illustrate the results on the average spectrum efficiency per user in Fig. 3.8. In the simulation, the average ISD is 500

meters, and $K = 10$, which is shown to be interference-limited in Fig. 3.1. Consistent with the SIR results, in a interference-limited network, ZF receivers provide a higher spectrum efficiency per user. Numerical results also show that the average spectrum efficiency is sensitive to the fraction of the path loss compensation ϵ; the optimum ϵ for per user rate is generally around 0.5 for MRC receiver, and 0.2 for ZF receivers. In addition, I also observe that there is a minor performance gap in rate between ZF and MRC receivers under full channel compensation power control, as predicted by Proposition 3.5.1.

Last, I examine the cell throughput in a system operated at 2 GHz in Fig. 3.9. As an example, I consider an OFDM system, where the symbol time is 66.7 μs. I consider two cases with different mobilities as listed in Table 3.1; the coherence time T_{c} is computed as $T_{\mathrm{c}} = \frac{1}{4f_{\mathrm{D}}}$ [138], where f_{D} is the maximum doppler frequency. In this simulation, I assume the density of overall users to be 100 times the base station density, to simulate the case with large K. In Fig. 3.9(a), in the high mobility case, when $T_{\mathrm{c}} < M$, the optimal K^* for cell throughput is limited by the duration of T_{c}, and the optimal value generally is $K^* \leq \frac{T_{\mathrm{c}}}{2}$. In the example of low mobility case, when $T_{\mathrm{c}} > M$, the results in Fig. 3.9(b) show that the optimal K^* depends much on M: for ZF receivers, the cell throughput drops fast when $\frac{K}{M}$ approaches to 1, and the optimal K^* is around $\frac{M}{2}$ for maximum throughput; for MRC receivers, the cell throughput becomes saturated approximately when $K > \frac{M}{3}$. In addition, ZF receivers generally achieve better cell throughput than MRC receivers; the only exception is the case of $\frac{M}{K} \approx 1$, where the cell throughput of ZF receivers drops below that of

MRC. In addition, compared with the single user per cell case ($K = 1$), the results confirm that massive MIMO improves the cell throughput by serving multiple users simultaneously.

3.7 Conclusion

In this chapter, I proposed a stochastic geometry framework to analyze the spatial average SINR coverage and rate in massive MIMO networks. I applied the analysis and numerical results to draw several important system design insights about the SINR coverage and rate in uplink massive MIMO networks.

- The uplink massive MIMO networks can be interference-limited in urban marco cells (ISD=500 meters) with $M = 64$ antennas at base stations.
- With MRC receivers, the number of antennas M should scale super-linearly with the number of scheduled users per cell K as $(M + 1) \sim K^{\frac{\alpha}{2}(1-\epsilon)+\epsilon}$, to maintain the uplink SIR distribution; a linear scaling law only applies to the case of full path loss compensation in the power control, i.e., when $\epsilon = 1$.
- With ZF receivers, a linear scaling between the number of antennas M and users per cell K maintains the uplink SIR distribution in massive MIMO.
- When noise is negligible, ZF receivers provide better SIR coverage rate

than MRC receivers. The performance gap increases with K, and decreases with path loss compensation faction ϵ. The gap becomes minor when $\epsilon = 1$.

- The SIR coverage and rate are sensitive to the fraction ϵ of path loss compensation in power control. Larger ϵ improves coverage in the low SIR regime while reducing coverage probability at high SIR. Numerical results show that the optimal ϵ for rate is around 0.5 for MRC, and 0.2 for ZF receivers in certain cases.

3.8 Appendix

3.8.1 Derivation of (3.10)

With the combining vector in (3.9), the SINR expression equals (3.20) as

$$\frac{P_0^{(1)} \left| \mathbf{u}_{00}^{(1)*} \bar{\mathbf{h}}_{00}^{(1)} \right|^2}{P_0^{(1)} \mathbb{E} \left| \mathbf{u}_{00}^{(1)*} \hat{\mathbf{h}}_{00}^{(1)} \right|^2 + \sum_{(\ell,k)\neq(0,1)} P_\ell^{(k)} \mathbb{E} \left| \mathbf{n}_{\mathrm{t}}^* \mathbf{h}_{0\ell}^{(k)} + \sum_{\ell' \geq 0} \sqrt{P_{\ell'}^{(1)}} \mathbf{h}_{0\ell'}^{(1)*} \mathbf{h}_{0\ell}^{(k)} \right|^2 + |\mathbf{u}_{00}^{(1)}|^2 \sigma^2}. \tag{3.20}$$

In the numerator, the signal power can be computed as

$$\begin{aligned} P_0^{(1)} |\mathbf{u}_{00}^{(1)*} \bar{\mathbf{h}}_{00}^{(1)}|^2 &\overset{(a)}{=} \frac{\left(P_0^{(1)} \beta_{00}^{(1)} \right)^2}{\left(\sum_\ell P_\ell^{(1)} \beta_{0\ell}^{(1)} + \frac{\sigma^2}{K} \right)^2} |\mathbf{u}_{00}^{(1)}|^4 \overset{(b)}{\approx} \frac{\left(P_0^{(1)} \beta_{00}^{(1)} \right)^2}{\left(\sum_\ell P_\ell^{(1)} \beta_{0\ell}^{(1)} + \frac{\sigma^2}{K} \right)^2} \mathbb{E} |\mathbf{u}_{00}^{(1)}|^4 \\ &\overset{(c)}{=} \left(P_0^{(1)} \beta_{00}^{(1)} \right)^2 (M^2 + M) = P_{\mathrm{t}}^2 \left(\beta_{00}^{(1)} \right)^{2(1-\epsilon)} (M^2 + M), \end{aligned}$$

where (a) follows from the MMSE estimator in (3.5), (b) follows from the fact that $|\mathbf{u}_{00}^{(1)}|^4 \overset{M\to\infty}{\to} \mathbb{E}|\mathbf{u}_{00}^{(1)}|^4$, and the approximation error decays as $\frac{1}{M^2}$[105], and

(c) follows from the fact that $\mathbb{E}|\mathbf{u}_{00}^{(1)}|^4 = (M^2+M)\left(\sum_\ell P_\ell^{(1)}\beta_{0\ell}^{(1)} + \frac{\sigma^2}{K}\right)^2$. Next, I compute the first term in the denominator of (3.20) as

$$\begin{aligned}
P_0^{(1)}\mathbb{E}|\mathbf{u}_{00}^{(1)*}\hat{\mathbf{h}}_{00}^{(1)}|^2 &= \frac{1}{M}P_0^{(1)}\mathbb{E}|\mathbf{u}_{00}^{(1)}|^2\mathbb{E}|\hat{\mathbf{h}}_{00}^{(1)}|^2 \\
&\overset{(a)}{=} \frac{1}{M}P_0^{(1)}M\left(\sum_\ell P_\ell^{(1)}\beta_{0\ell}^{(1)} + \frac{\sigma^2}{K}\right)M\beta_{00}^{(1)}\left(1 - \frac{P_0^{(1)}\beta_{00}^{(1)}}{\sum_\ell P_\ell^{(1)}\beta_{0\ell}^{(1)} + \frac{\sigma^2}{K}}\right) \\
&= MP_{\mathrm{t}}^2\left(\beta_{00}^{(1)}\right)^{1-\epsilon}\left(\sum_{\ell>0}\left(\beta_{\ell\ell}^{(1)}\right)^{-\epsilon}\beta_{0\ell}^{(1)} + \frac{\sigma^2}{P_{\mathrm{t}}K}\right),
\end{aligned}$$

where (a) follows from the fact that the channel estimation error $\hat{\mathbf{h}}_{00}^{(1)}$ follows the distribution $\mathcal{CN}\left(\mathbf{0}, \beta_{00}^{(1)}\left(1 - \frac{P_0^{(1)}\beta_{00}^{(1)}}{\sum_\ell P_\ell^{(1)}\beta_{0\ell}^{(1)}+\frac{\sigma^2}{K}}\right)\mathbf{I}_M\right)$.

Next, I simplify the second term in the denominator. Note that unless $(\ell,\ell',k) = (n,n',m)$, $\mathbf{h}_{\ell'0}^{(1)*}\mathbf{h}_{\ell 0}^{(k)}$ and $\mathbf{h}_{n'0}^{(1)*}\mathbf{h}_{n0}^{(m)}$ are uncorrelated zero-mean random variables. Therefore, I can simplify the second term in the denominator of (3.20) as

$$\sum_{(\ell,k)\neq(0,1)}\mathbb{E}|\mathbf{n}_{\mathrm{t}}^*\mathbf{h}_{\ell 0}^{(k)} + \sum_{\ell'\geq 0}\sqrt{P_{\ell'}^{(1)}}\mathbf{h}_{\ell'0}^{(1)*}\mathbf{h}_{\ell 0}^{(k)}|^2 \tag{3.21}$$

$$= \sum_{(\ell,k)\neq(0,1)}\left(\frac{\sigma^2}{K}MP_{\mathrm{t}}\left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon}\beta_{\ell 0}^{(k)} + \sum_{\ell'>0}P_\ell^{(k)}P_{\ell'}^{(1)}\mathbb{E}\left[|\mathbf{h}_{\ell'0}^{(1)*}\mathbf{h}_{\ell 0}^{(k)}|^2\right]\right). \tag{3.22}$$

Note that for $k=1$,$\ell' = \ell \neq 0$, the expression is simplified as

$$\begin{aligned}
P_\ell^{(k)}P_{\ell'}^{(1)}\mathbb{E}\left[|\mathbf{h}_{\ell'0}^{(1)*}\mathbf{h}_{\ell 0}^{(k)}|^2\right] &= P_\ell^{(1)2}\mathbb{E}\left[|\mathbf{h}_{\ell 0}^{(1)}|^4\right] \\
&= (M^2+M)P_\ell^{(1)2}\left(\beta_{0\ell}^{(1)}\right)^2 \\
&= P_{\mathrm{t}}^2(M^2+M)\left(\beta_{\ell\ell}^{(1)}\right)^{-2\epsilon}\left(\beta_{0\ell}^{(1)}\right)^2;
\end{aligned}$$

for $k > 1$ or $k = 1$, $\ell' \neq \ell > 0$, it follows that

$$P_\ell^{(k)} P_{\ell'}^{(1)} \mathbb{E}\left[|\mathbf{h}_{\ell' 0}^{(1)*} \mathbf{h}_{\ell 0}^{(k)}|^2\right] = M P_{\ell'}^{(1)} P_\ell^{(k)} \beta_{0\ell'}^{(1)} \beta_{0\ell}^{(k)}$$
$$= M P_{\mathrm{t}}^2 \left(\beta_{\ell'\ell'}^{(1)} \beta_{\ell\ell}^{(k)}\right)^{-\epsilon} \beta_{0\ell'}^{(1)} \beta_{0\ell}^{(k)}.$$

Therefore, I can express (3.22) as

$$\sum_{(\ell,k)\neq(0,1)} \left(\frac{\sigma^2}{K} M P_{\mathrm{t}} \left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon} \beta_{\ell 0}^{(k)} + \sum_{\ell'>0} P_\ell^{(k)} P_{\ell'}^{(1)} \mathbb{E}\left[|\mathbf{h}_{\ell' 0}^{(1)*} \mathbf{h}_{\ell 0}^{(k)}|^2\right]\right)$$
$$= M P_{\mathrm{t}}^2 \left(\sum_{(k,\ell)\neq(1,0)} \left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon} \beta_{\ell 0}^{(k)}\right) \left(\frac{\sigma^2}{K P_{\mathrm{t}}} + \sum_{\ell'\geq 0} \left(\beta_{\ell'\ell'}^{(1)}\right)^{-\epsilon} \beta_{0\ell'}^{(1)}\right)$$
$$+ M^2 P_{\mathrm{t}}^2 \sum_{\ell>0} \left(\beta_{\ell\ell}^{(1)}\right)^{-2\epsilon} \left(\beta_{0\ell}^{(1)}\right)^2.$$

Last, the thermal noise term in the denominator can be simplified as

$$|\mathbf{u}_{00}^{(1)}|^2 \sigma^2 = P_{\mathrm{t}} M \sigma^2 \left(\sum_{\ell} \left(\beta_{\ell\ell}^{(1)}\right)^{-\epsilon} \beta_{0\ell}^{(1)} + \frac{\sigma^2}{P_{\mathrm{t}} K}\right).$$

Then the expression in (3.10) is obtained through algebraic manipulation in the denominator.

3.8.2 Proof of Theorem 3.4.1

To allow for tractable computation and decouple the correlated terms in the denominator of (3.10), I propose the following approximations on the out-of-cell interference terms: for $k \in [1, K]$,

$$\Delta_1^{(k)} \overset{(a)}{\approx} \mathbb{E}\left(\sum_{\ell>0} \left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon} \beta_{0\ell}^{(k)}\right) + \frac{\sigma^2}{K P_{\mathrm{t}}} = \mathbb{E}\left(\sum_{\ell>0} \mathbb{E}\left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon} \beta_{0\ell}^{(k)}\right) + \frac{\sigma^2}{K P_{\mathrm{t}}}$$

$$\overset{(b)}{\approx} C^{-\epsilon}(\lambda_{\mathrm{b}}\pi)^{-\frac{\alpha\epsilon}{2}}\Gamma^{\alpha}(\frac{\epsilon}{2}+1)\mathbb{E}\left(\sum_{\ell>0}\beta_{0\ell}^{(k)}\right)+\frac{\sigma^2}{KP_{\mathrm{t}}}$$

$$\overset{(c)}{=} \frac{2C^{1-\epsilon}(\lambda_{\mathrm{b}}\pi)^{\frac{\alpha(1-\epsilon)}{2}}}{\alpha-2}\Gamma^{\alpha}\left(\frac{\epsilon}{2}+1\right)+\frac{\sigma^2}{KP_{\mathrm{t}}}, \tag{3.23}$$

where in (a), I approximate $\Delta_1^{(k)}$ by its mean; step (b) follows from

$$\mathbb{E}\left(\beta_{\ell\ell}^{(k)}\right)^{-\epsilon}=C^{-\epsilon}\left(\mathbb{E}\left[R_{\ell\ell}^{(k)\epsilon}\right]\right)^{\alpha}=(\lambda_{\mathrm{b}}\pi)^{-\frac{\alpha\epsilon}{2}}\Gamma^{\alpha}\left(\frac{\epsilon}{2}+1\right); \tag{3.24}$$

and step (c) follows from the exclusion ball model in Approximation 3.3.1 and the Campbell's theorem [24] as

$$\mathbb{E}\left[\sum_{\ell>0}\beta_{0\ell}^{(k)}\right]=2\pi\lambda_{\mathrm{b}}C\int_{R_{\mathrm{e}}}^{\infty}x^{-\alpha}x\mathrm{d}x=\frac{2C(\lambda_{\mathrm{b}}\pi)^{\alpha/2}}{\alpha-2}. \tag{3.25}$$

Similarly, I can approximate $\Delta_2^{(1)}$ by its mean as:

$$\Delta_2^{(k)}\approx\frac{C^{2(1-\epsilon)}\left(\lambda_{\mathrm{b}}\pi\right)^{\alpha(1-\epsilon)}}{\alpha-1}\Gamma^{\alpha}(\epsilon+1), \tag{3.26}$$

where follows from $\mathbb{E}\left(\beta_{\ell\ell}^{(k)}\right)^{-2\epsilon}=C^{-2\epsilon}\left(\mathbb{E}\left[R_{\ell\ell}^{(k)2\epsilon}\right]\right)^{\alpha}=(\lambda_{\mathrm{b}}\pi)^{-\alpha\epsilon}\Gamma^{\alpha}(\epsilon+1)$.

Next, applying the approximation in (3.23) and (3.26) and conditioning on $R_{00}^{(1)}=x$, I simplify the approximate conditional uplink SINR as

$$\mathbb{P}\left(\mathrm{SIR}>T|R_{00}^{(1)}=x\right)$$

$$\overset{(a)}{\approx}\mathbb{P}\left(1>\frac{T}{M+1}\left(C_1(\lambda_{\mathrm{b}}\pi x^2)^{\alpha(1-\epsilon)}+C_2(\lambda_{\mathrm{b}}\pi x^2)^{\frac{\alpha}{2}(1-\epsilon)}+\right.\right.$$

$$\left.\left.C_4(\lambda_{\mathrm{b}}\pi x^2)\sum_{k>0}\left(\lambda_{\mathrm{b}}\pi R_{00}^{(k)2}\right)^{-\frac{\alpha}{2}(1-\epsilon)}\right)\right)$$

$$\overset{(b)}{\approx}\mathbb{P}\left(g>\frac{T}{M+1}\left(C_1(\lambda_{\mathrm{b}}\pi x^2)^{\alpha(1-\epsilon)}+C_2(\lambda_{\mathrm{b}}\pi x^2)^{\frac{\alpha}{2}(1-\epsilon)}+\right.\right.$$

$$\left.\left.C_4(\lambda_{\mathrm{b}}\pi x^2)\sum_{k>0}\left(\lambda_{\mathrm{b}}\pi R_{00}^{(k)2}\right)^{-\frac{\alpha}{2}(1-\epsilon)}\right)\right)$$

$$
\begin{aligned}
&\overset{(c)}{\approx} 1-\mathbb{E}\left[\left(1-\exp\left(-\frac{\eta T}{M+1}\left(C_1(\lambda_{\mathrm{b}}\pi x^2)^{\alpha(1-\epsilon)}+\right.\right.\right.\right.\\
&\left.\left.\left.\left.C_2(\lambda_{\mathrm{b}}\pi x^2)^{\frac{\alpha}{2}(1-\epsilon)}+C_4(\lambda_{\mathrm{b}}\pi x^2)\sum_{k>0}\left(\lambda_{\mathrm{b}}\pi R_{00}^{(k)2}\right)^{-\frac{\alpha}{2}(1-\epsilon)}\right)\right)\right)^N\right]\\
&=\sum_{n=1}^{N}\binom{N}{n}(-1)^{n+1}\mathrm{e}^{-nC_0\left(C_1\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\alpha(1-\epsilon)}+C_2\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\frac{\alpha}{2}(1-\epsilon)}\right)}\times\\
&\mathbb{E}\left[\mathrm{e}^{-nC_0C_4\left(\lambda_{\mathrm{b}}\pi x^2\right)\left(\lambda_{\mathrm{b}}\pi R_{00}^{(k)2}\right)^{-\frac{\alpha}{2}(1-\epsilon)}}\right]^{K-1}\\
&\overset{(d)}{=}\sum_{n=1}^{N}\binom{N}{n}(-1)^{n+1}\mathrm{e}^{-nC_0\left(C_1\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\alpha(1-\epsilon)}+C_2\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\frac{\alpha}{2}(1-\epsilon)}\right)}\times\\
&\left(\int_0^{\infty}\mathrm{e}^{-nC_0C_4\left(\lambda_{\mathrm{b}}\pi x^2\right)s^{-\alpha(1-\epsilon)}-\lambda_{\mathrm{b}}\pi s^2}\lambda_{\mathrm{b}}\pi s\mathrm{d}s\right)^{K-1}\\
&\overset{(e)}{=}\sum_{n=1}^{N}\binom{N}{n}(-1)^{n+1}\mathrm{e}^{-nC_0\left(C_1\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\alpha(1-\epsilon)}+C_2\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\frac{\alpha}{2}(1-\epsilon)}\right)}\\
&\left(\int_0^{\infty}\mathrm{e}^{-nC_0C_4\left(\lambda_{\mathrm{b}}\pi x^2\right)u^{-\frac{\alpha}{2}(1-\epsilon)}-u}\mathrm{d}u\right)^{K-1}\times\\
&\overset{(f)}{\approx}\sum_{n=1}^{N}\binom{N}{n}(-1)^{n+1}\mathrm{e}^{-nC_0\left(C_1\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\alpha(1-\epsilon)}+C_2\left(\lambda_{\mathrm{b}}\pi x^2\right)^{\frac{\alpha}{2}(1-\epsilon)}\right)}\times\\
&\left(1-nC_0C_4\left(\lambda_{\mathrm{b}}\pi x^2\right)\int_0^{\infty}\frac{\mathrm{e}^{-u}\mathrm{d}u}{1+nC_0C_4\left(\lambda_{\mathrm{b}}\pi x^2\right)u^{-\frac{\alpha}{2}(1-\epsilon)}}\right)^{K-1},
\end{aligned}
$$

where in (a) C_1, C_2 and $C_4(x)$ are defined in Theorem 3.4.1; in (b) I use a "dummy" gamma variable g with unit mean and shape parameter N to approximate the constant number one, and the approximation follows from the fact that g converges to one when N goes to infinity, i.e., $\lim_{n\to\infty}\frac{n^n x^{n-1}\mathrm{e}^{-nx}}{\Gamma(n)}=\delta(x-1)$ [20], where $\delta(x)$ is the Dirac delta function; in (c), the approximation follows from Alzer's inequality [14, 29, Appendix A], where $\eta=N(N!)^{-\frac{1}{N}}$; (d) follows from the fact that $R_{00}^{(k)}$ is assumed to be IID Rayleigh random variable; in (e) I change variable as $u=\lambda_{\mathrm{b}}\pi s^2$; in (f), I apply the approximation

$\exp(-x) \approx \frac{1}{1+x}$ inside the integral, to allow for faster numerical evaluations. Last, I obtain the uplink SIR distribution by de-conditioning on $R_{00}^{(1)} = x$, which is assumed to be a Rayleigh random variable with mean $0.5\sqrt{1/\lambda_{\rm b}}$, and changing the variable as $t = \pi\lambda_{\rm b}x^2$.

3.8.3 Proof of Theorem 3.5.1

Before proving the theorem, I present a useful lemma on the distribution of the combining vector $\mathbf{g}_{00}^{(1)}$ as follows.

Lemma 3.8.1 (From [71]). *The square norm of the combining vector* $|\mathbf{g}_{00}^{(1)}|^2$ *follows the distribution of* $\left(\chi^2_{2(K-M+1)}\left(\sum_{\ell'} P_{\ell'}^{(1)}\beta_{\ell\ell'}^{(1)} + \frac{\sigma^2}{K}\right)\right)^{-1}$, *where* $\chi^2_{2(K-M+1)}$ *represents a Chi-square random variables with* $2(K-M+1)$ *degrees of freedom.*

Note that when $(M_K + 1) \to \infty$, $\frac{\chi^2_{2(K-M+1)}}{M-K+1} \to 1$. Therefore, by Lemma 3.8.1, when $M \gg K$, the following approximation holds as

$$|\mathbf{g}_{00}^{(1)}|^2 \approx \frac{1}{(M-K+1)S^{(1)}}, \tag{3.27}$$

where for ease of notation, I define $S^{(k)} = \sum_{\ell'} P_{\ell'}^{(1)}\beta_{\ell\ell'}^{(1)} + \frac{\sigma^2}{K} = \left(\beta_{00}^{(k)}\right)^{1-\epsilon} + \Delta_1^{(k)}$, and the approximation error decays as $\frac{1}{M-K+1}$. Noting that $|\mathbf{g}_{00}^{(1)}\mathbf{u}_{00}^{(1)}|^2 = 1$, the nominator of the SINR expression in (3.8) can be computed as

$$P_0^{(1)}|\mathbf{g}_{00}^{(1)*}\bar{\mathbf{h}}_{00}^{(1)}|^2 = \frac{P_0^{(1)2}\beta_{00}^{(1)2}}{\left(S^{(1)}\right)^2}|\mathbf{g}_{00}^{(1)}\mathbf{u}_{00}^{(1)}|^2 = \left(\beta_{00}^{(1)}\right)^{2(1-\epsilon)} / \left(S^{(1)}\right)^2. \tag{3.28}$$

Applying the results in (3.27), the first term in the denominator of (3.8)

is computed as

$$P_0^{(1)}\mathbb{E}|\mathbf{g}_{00}^{(1)*}\hat{\mathbf{h}}_{00}^{(1)}|^2 \approx P_0^{(1)}\beta_{00}^{(1)}\left(1-\frac{P_0^{(1)}\beta_{00}^{(1)}}{S^{(1)}}\right)\frac{1}{(M-K+1)S^{(1)}}$$
$$=\frac{\Delta_1^{(1)}\left(\beta_{00}^{(1)}\right)^{1-\epsilon}}{(M-K+1)\left(S^{(1)}\right)^2}. \tag{3.29}$$

Now I simplify the second sum in the denominator of (3.8): for $k=1$, and $\ell>0$, it follows

$$\mathbb{E}|\mathbf{g}_{00}^{(1)*}\mathbf{h}_{0\ell}^{(1)}|^2=\mathbb{E}|\mathbf{u}_{00}^{(1)}|^{-4}|\mathbf{g}_{00}^{(1)*}\mathbf{u}_{00}^{(1)}\mathbf{u}_{00}^{(1)*}\mathbf{h}_{\ell 0}^{(1)}|^2=\mathbb{E}|\mathbf{u}_{00}^{(1)}|^{-4}|\mathbf{u}_{00}^{(1)*}\mathbf{h}_{\ell 0}^{(1)}|^2$$
$$=\frac{(M^2+M)P_\ell^{(1)2}\beta_{0\ell}^{(1)2}+M\sum_{\ell'\neq\ell}P_\ell^{(1)}P_{\ell'}^{(1)}\beta_{0\ell}^{(1)}\beta_{0\ell'}^{(1)}+\frac{M\sigma^2}{K}P_\ell^{(1)}\beta_{0\ell}^{(1)}}{(M^2+M)\left(S^{(1)}\right)^2}; \tag{3.30}$$

For $k>1$, it follows that

$$\mathbb{E}|\mathbf{g}_{00}^{(1)*}\mathbf{h}_{0\ell}^{(k)}|^2=\frac{\mathbb{E}|\mathbf{g}_{00}^{(1)}|^2\mathbb{E}\left(|\mathbf{h}_{0\ell}^{(k)}|^2-|\mathbf{u}_{00}^{(k)}|^{-2}|\mathbf{u}_{00}^{(k)*}\mathbf{h}_{\ell 0}^{(k)}|^2\right)}{M-K+1}$$
$$=\frac{\mathbb{E}\left(|\mathbf{h}_{0\ell}^{(k)}|^2-|\mathbf{u}_{00}^{(k)}|^{-2}|\mathbf{u}_{00}^{(k)*}\mathbf{h}_{\ell 0}^{(k)}|^2\right)}{(M-K+1)^2S^{(1)}}\approx\frac{MP_\ell^{(k)}\beta_{0\ell}^{(k)}\left(1-\frac{P_\ell^{(k)}\beta_{0\ell}^{(k)}}{S^{(k)}}\right)}{(M-K+1)^2S^{(1)}}.$$

Then, in the case of $k>1$, for $\ell=0$, I approximate the residue intra-cell interference of ZF receivers as

$$\mathbb{E}|\mathbf{g}_{00}^{(1)*}\mathbf{h}_{00}^{(k)}|^2\approx\frac{MC^{1-\epsilon}(\lambda_{\mathrm{b}}\pi)^{\frac{\alpha(1-\epsilon)}{2}}\Delta_1^{(k)}}{(M-K+1)^2\left(C^{1-\epsilon}(\lambda_{\mathrm{b}}\pi)^{\frac{\alpha(1-\epsilon)}{2}}+\Delta_1^{(k)}\right)S^{(1)}}, \tag{3.31}$$

which follows from $P_0^{(k)}\beta_{00}^{(k)}\approx\left(\mathbb{E}\left[R_{00}^{(k)2}\right]\right)^{-\frac{\alpha(1-\epsilon)}{2}}=(\lambda_{\mathrm{b}}\pi)^{\frac{\alpha(1-\epsilon)}{2}}$; for $\ell>0$, the out-of-cell interference is upper bounded (and approximated) as

$$\mathbb{E}|\mathbf{g}_{00}^{(1)*}\mathbf{h}_{0\ell}^{(k)}|^2\leq\frac{MP_1\beta_{0\ell}^{(k)}}{(M-K+1)^2S^{(1)}}. \tag{3.32}$$

The noise term in the denominator is

$$|\mathbf{g}_{00}^{(1)}|^2\sigma^2 = \frac{\sigma^2}{(M-K+1)S^{(1)}}. \tag{3.33}$$

Next, substituting (3.23), (3.26) and (3.29)-(3.33) for (3.8), conditioning on $R_{00}^{(1)} = x$, and after some algebraic manipulation, the SINR expression is simplified as

$$\text{SINR} = \left(C_6(\lambda_\text{b}\pi x^2)^{\frac{\alpha}{2}(1-\epsilon)} + C_7(\lambda_\text{b}\pi x^2)^{\alpha(1-\epsilon)}\right)^{-1}.$$

The rest of the proof follows the same line as in Section 3.8.2.

Chapter 4

Analysis of Millimeter Wave Massive MIMO Networks

4.1 Motivation and Related Work

Massive MIMO is a promising candidate technology for 5G cellular networks [76, 42, 87]. It deploys more antennas than in conventional systems to serve a large number of users and provides high throughput [90, 42, 76, 87]. With large bandwidth channels, the mmWave spectrum constitutes a promising candidate frequency for access channels in 5G cellular networks [112, 120]. The small wavelength also makes it natural to consider massive MIMO at mmWave frequencies [133], which I call *mmWave massive MIMO* in this chapter. MmWave cellular networks will be different from the system at lower frequencies. One key difference is the sensitivity to blockages: different path loss laws are found in the LOS and NLOS mmWave links in measurements [120]. Besides the differences in propagation, mmWave base stations will probably have fewer RF chains than conventional systems, due to power constraints. Consequently, a mmWave base station will support fewer users, and probably use analog or hybrid beamforming [11].

Stochastic geometry provides useful tools to analyze the system-level

performance in large-scale networks [18]. In [89, 27], the asymptotic performance in a conventional massive MIMO network was examined using stochastic geometry, where key features of mmWave systems, e.g. the blockages and directional beamforming, were not incorporated. Stochastic geometric cellular models [18] were also extended to analyze mmWave network performance [29, 25]. An innovation of the analysis in [29, 25] was to incorporate building blockage into the analytical framework by using different path-loss laws for LOS and NLOS links. Unfortunately, the framework in [25, 29] cannot be directly applied to study mmWave massive MIMO networks, as intra-cell interference was not treated by assuming a single user per cell, and the channel training stage was not included by assuming perfect channel knowledge.

4.2 Contributions

In this chapter, I propose to study the asymptotic SINR and rate performance in a TDD mmWave massive MIMO system. I extend the mmWave cellular model in [29, 25] to the mmWave massive MIMO case by characterizing the distributions of the multiple scheduled users per cell and incorporating the difference in spatial correlations between LOS and NLOS links. Based on the system model, I derive expressions to evaluate the asymptotic SINR distributions in both uplink and downlink, when the number of base station antennas goes to infinity. Compared with prior asymptotic analysis of massive MIMO in [90], I incorporate key features in mmWave networks, including blockage effects and directional beamforming, and consider a different large-scale net-

work topology with infinite randomly distributed base stations. Numerical results show that the SINR performance in massive MIMO networks is dependent on the base station density, where a good SINR coverage and high convergence rate to the asymptotic equivalence are achieved with a dense base station deployment. Moreover, a comparison with massive MIMO system at 2 GHz shows that mmWave massive MIMO achieves a higher cell throughput when densely deployed.

4.3 System Model

In this section, I introduce the system model for a mmWave massive MIMO network. I consider a mmWave massive MIMO cellular network with perfect synchronization. Each base station is assumed to have M antennas. In each time-frequency resource block, a base station can schedule K users simultaneously in its cell. Let X_ℓ be the location of the ℓ-th base station, $Y_\ell^{(k)}$ be the location of the k-th scheduled user in the cell of ℓ-th base station, and $\mathbf{h}_{\ell\ell'}^{(k)}$ the channel vector from X_ℓ to $Y_{\ell'}^{(k)}$.

The network is assumed to be operated in the following TDD mode as proposed in [90]: channel training is performed in the uplink, where the users send their assigned pilots $\mathbf{T}_k$, and base stations estimate the channels using the orthogonality of the pilots; based on the channel estimates derived from uplink pilots, the base stations apply maximum ratio combining to receive the uplink data, and match-filter beamforming to transmit the downlink data. Further, I assume full reuse of the orthogonal pilots $\{\mathbf{T}_k\}_{1\leq k\leq K}$ in the network.

Due to the power and hardware constraints [11], mmWave base stations will have fewer RF chains than the antennas. To address the limitation on the number of RF chains, I restrict the number of simultaneously scheduled users in a cell to be smaller than 4. Though mmWave systems will probably apply analog or hybrid beamforming, for simplicity, I assume digital beamforming and combining in this chapter.

Now, I introduce the mmWave channel assumptions. One key feature of mmWave channels is the sensitivity to blockages: the presence of building blockages in urban areas makes the large-scale path loss laws much different in the LOS and NLOS links [120]. To distinguish the LOS and NLOS links, let $p_{\mathrm{L}}(R)$ be the probability that a link of length R is LOS, and $p_{\mathrm{N}}(R)$ the probability that the link is NLOS. The LOS probability function $p_{\mathrm{L}}(R)$ is assumed to have finite first moment, i.e., $\int_0^\infty p_{\mathrm{L}}(r) r \mathrm{d}r < \infty$, which is satisfied by most LOS probability functions in literatures, e.g. the ones in [30, 1]. The path loss $L(R)$ for a link of length R is

$$\begin{aligned} L(R) = {} & \mathbf{1}\left[\psi < p_{\mathrm{L}}(R)\right] C_{\mathrm{L}} (\max(\delta, R))^{-\alpha_{\mathrm{L}}} \\ & + \mathbf{1}\left[p_{\mathrm{L}}(R) \leq \psi < (p_{\mathrm{L}}(R) + p_{\mathrm{N}}(R))\right] C_{\mathrm{N}} (\max(\delta, R))^{-\alpha_{\mathrm{N}}}, \end{aligned} \tag{4.1}$$

where $\mathbf{1}[\cdot]$ is the indicator function, ψ is a uniform random variable in $[0, 1]$, $\delta = 1$ meter is the reference distance, α_{L}, α_{N} are the LOS and NLOS path loss exponents, and C_{L}, C_{N} are the intercepts in the LOS and NLOS path loss. Typical values of mmWave path loss parameters are available in measurement chapters [6, 120]. In addition, I assume for different links, the random variable

ψ, which decides whether the link is LOS or not, are independent.

For small-scale fading, I consider narrowband channels, as the effect of frequency selective fading can be minimized by techniques like orthogonal frequency-division multiplexing (OFDM) or frequency domain equalization [59]. Measurements show that small-scale fading has minor effects on LOS mmWave signals [120]. Consequently, for LOS links, I model the channel vector $\mathbf{h}_{\ell\ell'}^{(k)}$ by a deterministic vector as

$$\mathbf{h}_{\ell\ell'}^{(k)} = \sqrt{M}\beta_{\ell\ell'}^{(k)1/2}\mathbf{u}_{\ell\ell'}^{(k)}, \tag{4.2}$$

where $\beta_{\ell\ell'}^{(k)}$ is the path loss computed from (4.1), assuming the link is LOS, i.e., $\psi < p_{\mathrm{L}}(R)$; $\mathbf{u}_{\ell\ell'}^{(k)}$ is a unit vector. Furthermore, I assume asymptotic orthogonality between any two LOS channels as for $(s,t,k) \neq (s',t',k')$,

$$\lim_{M\to\infty} \mathbf{u}_{st}^{(k)*}\mathbf{u}_{s't'}^{(k')} = 0. \tag{4.3}$$

One example satisfying (4.3) is the LOS channels using uniform linear arrays with non-overlapping angles of arrival [105].

For the NLOS channel, I apply a correlated fading model to account for the potentially larger number of scatters. I express the NLOS channel vector as

$$\mathbf{h}_{\ell n}^{(k)} = \left(\beta_{\ell n}^{(k)}\right)^{1/2} \mathbf{\Phi}_{\ell n}^{(k)1/2}\mathbf{w}_{\ell n}^{(k)}, \tag{4.4}$$

where $\beta_{\ell n}^{(k)}$ is the large-scale path loss, $\mathbf{w}_{\ell n}^{(k)}$ is the fading vector consisting of IID random variables with zero mean and unit variance, and $\mathbf{\Phi}_{\ell n}^{(k)}$ is the

covariance matrix to account for correlations in small-scale fading. Let $\lambda_{\ell n}^{(k)}[m]$ be the eigenvalues of the covariance matrix $\mathbf{\Phi}_{\ell n}^{(k)}$. I assume the traces of the covariance matrices for all channels are normalized to M, i.e., Trace $\left[\mathbf{\Phi}_{\ell n}^{(k)}\right] = \sum_{m=1}^{M} \lambda_{\ell n}^{(k)}[m] = M$, and the average squares of the eigenvalues are upper bounded by a constant γ: $\frac{1}{M}\sum_{m=1}^{M} \lambda_{\ell n}^{(k)2}[m] \leq \gamma$, which is satisfied by many common channel models, including the case of uniform linear arrays with half-wavelength spacing and continuous angular spread [4].

Next, I describe the stochastic geometric network model. I assume the base stations are distributed as a PPP with density λ_{b}. The users, either scheduled or not, are distributed as an independent PPP on the plane with sufficiently high density, such that each base station is associated with at least K users. A user is assumed to be associated with the base station that provides the minimum path loss signal. Without loss of generality, a typical scheduled user $Y_0^{(1)}$ is fixed at the origin. I will investigate the SINR and rate performance at this typical user.

Now I focus on the distribution of scheduled users. In a resource block, let $\mathcal{N}_{\mathrm{u}}^{(k)}$ be the point process formed by the locations of the scheduled users $Y_{\ell}^{(k)}$, i.e., all the scheduled users assigned with the k-th pilot sequence. Note that though the users are distributed as a PPP on the plane, the scheduled users do not form a PPP, as their locations are correlated. For instance, in each cell, the number of scheduled users is fixed to be K, while for a PPP, the number of points in the same region is randomly distributed. The correlations in the scheduled users' location make the exact analysis intractable. Therefore,

I make the following approximation on the distribution of $\mathcal{N}_{\mathrm{u}}^{(k)}$.

Approximation 4.3.1. *The path losses $\beta_{\ell\ell}^{(k)}$ from the users to their associated base stations are assumed to be IID. Moreover, the tagged base station X_0 is assumed to observe the other-cell scheduled users in $\mathcal{N}_{\mathrm{u}}^{(k)}$ as a thinned Poisson point process with intensity function*

$$\lambda_{\mathrm{u}}(|X_0 - x|) = \lambda_{\mathrm{b}} \mathbf{1}(L(|X_0 - x|) > \beta_x), \tag{4.5}$$

where $L(|X_0 - x|)$ represents the path loss from x to X_0, β_x is an IID random variable with the same distribution as $\beta_{00}^{(1)}$, and the indicator function $\mathbf{1}(L(|X_0 - x|) > \beta_x)$ ensures that any user outside the tagged cell X_0 has smaller path loss to its own base station than to X_0. Besides, the scheduled users assigned with different pilots are assumed to be independently distributed, i.e., for $k \neq k'$, $N_{\mathrm{u}}^{(k)}$ and $N_{\mathrm{u}}^{(k')}$ are independent.

It can be shown in simulations that the proposed approximation provides a close characterization of the actual scheduled users in a network with PPP distributed base stations [33].

MmWave handsets will use antenna arrays to perform directional beamforming [112]. To simplify the analysis, the antenna array at the mobile station is modeled as a single directional antenna with a gain pattern approximated by the sectored antenna pattern. In the sectored antenna model, the directivity gain within the main lobe θ is assumed to be a constant Q, while all angles outside the main lobe θ have the constant side lobe gain q. Let $D_{\ell\ell'}^{(k)}$

be the directivity gain of the mobile station $Y_{\ell'}^{(k)}$ to base station X_ℓ. I assume the directions of the mobile station antennas are adjusted to maximize the desired signal power without alignment errors. Moreover, for the interfering links, I assume the antenna directions are independently and uniformly distributed. Consequently, the directivity gain $D_{\ell\ell}^{(k)} = Q$ in the desired link; for other interfering links, $D_{\ell\ell'}^{(k)}$ is a Bernoulli random variable, where $D_{\ell\ell'}^{(k)} = Q$ with probability $\frac{\theta}{2\pi}$, and $D_{\ell\ell'}^{(k)} = q$ otherwise.

Last, I assume that the base stations estimate the channels by correlating the received signals with the pilots, and do not use MMSE estimation. Hence, the channel estimate of $\mathbf{h}_{\ell\ell}^{(k)}$ at base station X_ℓ is

$$\bar{\mathbf{h}}_{\ell\ell}^{(k)} = \sqrt{Q}\mathbf{h}_{\ell\ell}^{(k)} + \sum_{\ell' \neq \ell} D_{\ell\ell'}^{(k)1/2}\mathbf{h}_{\ell\ell'}^{(k)} + \mathbf{n}_{\mathrm{T}},$$

where $\mathbf{n}_T$ is a Gaussian noise vector of the distribution $\mathcal{CN}\left(\mathbf{0}, \frac{1}{K\rho_{\mathrm{T}}}\mathbf{I}\right)$, and ρ_{T} is the SNR in the channel training.

Based on the proposed system model, I will analyze the SINR and rate performance in both uplink and downlink in the subsequent sections.

4.4 SINR Coverage Analysis

In this section, I derive analytical expressions to approximate the distributions of the asymptotic SINR, when the number of base station antennas goes to infinity.

4.4.1 Uplink Analysis

In this section, I analyze the uplink SINR performance in mmWave massive MIMO networks. First, I compute the density function of scheduled users in (4.5) explicitly as follows.

Lemma 4.4.1. *The density function* $\lambda_{\mathrm{u}}(r)$ *of the other-cell scheduled users in* $\mathcal{N}_{\mathrm{u}}^{(k)}$ *can be computed as*

$$\lambda_{\mathrm{u}}(r) = \lambda_{\mathrm{u,L}}(r) + \lambda_{\mathrm{u,N}}(r),$$

where for $\mathrm{s} \in \{\mathrm{L}, \mathrm{N}\}$, $\lambda_{\mathrm{u,s}}(r) = A\lambda_{\mathrm{b}} p_{\mathrm{s}}(r) \left(1 - \mathrm{e}^{-\Xi(r^{\alpha_{\mathrm{s}}}/C_{\mathrm{s}})}\right)$,

$$A = 1 - \mathrm{e}^{-2\pi\lambda_{\mathrm{b}} \int_0^\infty (p_{\mathrm{L}}(x) + p_{\mathrm{N}}(x)) x \mathrm{d}x},$$

and

$$\Xi(t) = 2\pi\lambda_{\mathrm{b}} \left(\int_0^{(tC_{\mathrm{N}})^{1/\alpha_N}} r p_{\mathrm{N}}(r) \mathrm{d}r + \int_0^{(tC_{\mathrm{L}})^{1/\alpha_L}} r p_{\mathrm{L}}(r) \mathrm{d}r \right).$$

Proof. See Section 4.7.1 □

To decode the uplink symbol $s_0^{(1)}$ sent by $Y_0^{(1)}$, the base station X_0 applies maximum ratio combining, based on the channel $\bar{\mathbf{h}}_{00}^{(1)}$. Then, the uplink SINR for the user $Y_0^{(1)}$ is

$$\frac{|\bar{\mathbf{h}}_{00}^{(1)*} \mathbf{h}_{00}^{(1)}|^2}{\sum_{\ell \neq 0} |\bar{\mathbf{h}}_{00}^{(1)*} \mathbf{h}_{0\ell}^{(1)}|^2 + \sum_{k=2}^{K} \sum_{\ell > 0} |\bar{\mathbf{h}}_{00}^{(1)*} \mathbf{h}_{0\ell}^{(k)}|^2 + |\bar{\mathbf{h}}_{00}^{(1)*} \mathbf{n}_{\mathrm{u}}|^2},$$

where $\mathbf{n}_{\mathrm{U}}$ is the noise vector of the distribution $\mathcal{CN}(\mathbf{0}, \rho_{\mathrm{U}}^{-1}\mathbf{I})$, and ρ_{U} is the SNR in uplink data transmission. When the number of antennas M goes to infinity, the uplink SINR converges to its asymptotic equivalence as follows.

Theorem 4.4.1. *When $M \to \infty$, the uplink SINR converges in probability to*

$$\mathrm{SINR}_{\mathrm{UL}} \xrightarrow{p.} \frac{Q^2 \beta_{00}^{(1)2}}{\sum_{\ell \neq 0} D_{0\ell}^{(1)2} \beta_{0\ell}^{(1)2}}. \tag{4.6}$$

The result in Theorem 4.4.1 is different from prior work in [90] in that (i) I consider a large-scale network with infinite randomly located base stations, in which case the analysis based on a network topology with finite base stations in [90] does not directly apply; (ii) I consider different spatial correlations in fading for the LOS and NLOS channels, while IID fading was assumed in [90]. Consequently, mathematical tools from stochastic geometry, including the factorial moment and Campbell's formula, are required to prove Theorem 4.4.1 [33]. The same comment also applies to Theorem 4.4.2 in the downlink analysis.

Next, I derive a tight approximation for the asymptotic SINR distribution in the following corollary.

Corollary 4.4.1.1. *The distribution of the asymptotic SINR in (4.6) can be approximated as*

$$\mathbb{P}\left(\mathrm{SINR}_{\mathrm{UL}} > T\right) \approx A \sum_{n=1}^{N} \binom{N}{n} (-1)^{n+1} \times \int_0^{\infty} \mathrm{e}^{-W_n(T,t) - V_n(T,t) - \Xi(t)} \Xi(\mathrm{d}t),$$

where

$$W_k(T,t) = \sum_{\ell=1}^{2} 2\pi B_\ell \int_0^{\infty} \left(1 - \mathrm{e}^{-k\eta T C_{\mathrm{L}} \xi_\ell t^2 x^{-2\alpha_L}}\right) \lambda_{\mathrm{u,L}}(x) x \mathrm{d}x,$$

$$V_k(T,t) = \sum_{\ell=1}^{2} 2\pi B_\ell \int_0^{\infty} \left(1 - \mathrm{e}^{-k\eta T C_{\mathrm{N}} \xi_\ell t^2 x^{-2\alpha_N}}\right) \lambda_{\mathrm{u,N}}(x) x \mathrm{d}x,$$

A is as defined in Lemma 4.4.1, N is the number of truncated terms used in the approximation, $\eta = N(N!)^{-\frac{1}{N}}$; for $\ell = 1, 2$, $B_\ell = \left\{\frac{\theta}{2\pi}, 1 - \frac{\theta}{2\pi}\right\}$, and $\xi_\ell = \{1, q^2/Q^2\}$.

Proof. See Section 4.7.2. □

Note that the expression in Corollary 4.4.1.1 can be further simplified when provided the explicit forms of $p_{\mathrm{L}}(r)$ and $p_{\mathrm{N}}(r)$, e.g. the step functions in [29, Section III-D]. Moreover, the expression becomes more accurate when more truncated terms (a larger N) is used; simulations indicate that using $N \geq 5$ terms is sufficient to provides a tight approximation.

4.4.2 Downlink SINR Analysis

In this section, I investigate the downlink asymptotic SINR in mmWave massive MIMO networks. Let $\mathbf{f}_\ell^{(k)}$ be the beamforming precoder that base station X_ℓ applies for its user $Y_\ell^{(k)}$. When applying match-filter beamforming, it follows that $\mathbf{f}_\ell^{(k)} = \frac{\bar{\mathbf{h}}_{\ell\ell}^{(k)}}{||\bar{\mathbf{h}}_{\ell\ell}^{(k)}||}$. Then the downlink SINR at the typical user $Y_0^{(1)}$ is

$$\frac{|\mathbf{h}_{00}^{(1)*}\mathbf{f}_0^{(1)}|^2}{\sum_{\ell\neq 0}|\mathbf{h}_{\ell 0}^{(1)*}\mathbf{f}_\ell^{(1)}|^2 + \sum_{k\neq 1}^{K}\sum_{\ell\geq 0}|\mathbf{h}_{\ell 0}^{(1)*}\mathbf{f}_\ell^{(k)}|^2 + \rho_{\mathrm{D}}^{-1}},$$

where ρ_{D} is the SNR in the downlink transmission. I have the following convergence result on the downlink SINR.

Theorem 4.4.2. *With match-filter beamforming, the downlink SINR will converge in probability to its asymptotic equivalence as*

$$\lim_{M\to\infty} \mathrm{SINR}_{\mathrm{DL}} \xrightarrow{p} \frac{Q^2\beta_{00}^{(1)2}/a_0^{(1)}}{\sum_{\ell\neq 0} D_{\ell 0}^{(1)2}\beta_{\ell 0}^{(1)2}/a_\ell^{(1)}},$$

where $a_\ell^{(k)} = \frac{1}{K\rho_{\mathrm{T}}} + \sum_{\ell'} D_{\ell\ell'}^{(k)}\beta_{\ell\ell'}^{(k)}$.

The exact distribution of the asymptotic SINR in Theorem 4.4.2 is generally difficult to derive, as the normalization constants $a_\ell^{(k)}$ introduce correlations among all the terms. In a dense mmWave network, however, the asymptotic SINR can be approximated by the following corollary.

Corollary 4.4.2.1. *In a dense mmWave network, the asymptotic downlink SINR distribution can be approximated as*

$$\mathbb{P}\left(\mathrm{SINR} > T\right) \approx A\sum_{n=1}^{N}\binom{N}{n}(-1)^{n+1}\int_0^\infty \mathrm{e}^{-Z_n(T,t)-\Xi(t)}\Xi(\mathrm{d}t),$$

where $Z_k(T,t) = \sum_{\ell=1}^{2} B_\ell \int_t^\infty \left(1-\mathrm{e}^{-k\eta T\xi_\ell t^2 x^{-2}}\right)\Xi(\mathrm{d}x)$, *$A$, η, B_ℓ, and ξ_ℓ are the same as defined in Corollary 4.4.1.1, and N is the number of terms used in the approximation.*

Proof. The proof is omitted here, as it is similar to that in the uplink in Section 4.7.2. □

Last, I define the average achievable rate at a typical user as $\Gamma = W(1-\mu)\log_2\left(1+\min\{\mathrm{SINR}, T_{\max}\}\right)$, where W is the bandwidth assigned to a user, μ is the fraction of overhead, and $T_{\max}$ is a SINR distortion threshold. In an OFDM massive MIMO system, the fraction of overhead μ can be computed

based on [90, Section III-A], which accounts for the inefficiency from cyclic prefix and uplink training. The use of a distortion threshold $T_{\max}$ is needed because of the potential for very high SINRs in massive MIMO networks that may not be exploited due to other limiting factors like linearity in the radio frequency front-end. By [18], given the SINR distribution $\mathbb{P}(\text{SINR} > T)$, the average achievable rate can be computed as $\mathbb{E}[\Gamma] = \frac{W}{\ln 2} \int_0^{T_{\max}} \frac{\mathbb{P}(\text{SINR}>t)}{1+t} \mathrm{d}t$.

4.5 Numerical Results

In this section, I consider a 28 GHz massive MIMO network with a system bandwidth of 500 MHz. I assume the transmitting power is 30 dBm in the downlink, and 20 dBm in the uplink. In the Monte Carlo simulations, the density of the user process is assumed to be 60 times the base station density, and a mmWave base station randomly serves $K = 4$ users in its cell in a resource block. Based on the New York city measurement in [6], I assume $p_{\mathrm{L}}(r) = \mathrm{e}^{-\xi r}\mathbf{1}(r < R_{\mathrm{O}})$ and $p_{\mathrm{N}}(r) = \left(1 - \mathrm{e}^{-\xi r}\right)\mathbf{1}(r < R_{\mathrm{O}})$, where $\xi = 70$ meters, and $R_{\mathrm{O}} = 200$ meters. In addition, I assume the LOS path loss exponent is $\alpha_{\mathrm{L}} = 2$, and the NLOS path loss exponent is $\alpha_{\mathrm{N}} = 4$. For the channel models, I use the steering vectors of uniform linear array with half-wavelength spacing as LOS channel vectors; for NLOS channels, I assume IID Rayleigh fading for simplicity.

First, I show the convergence of uplink SINR in a dense mmWave network in Fig. 4.1. In the simulations, I assume each base station schedule $K = 4$ users, and mobile stations use omni-directional antennas. As shown in

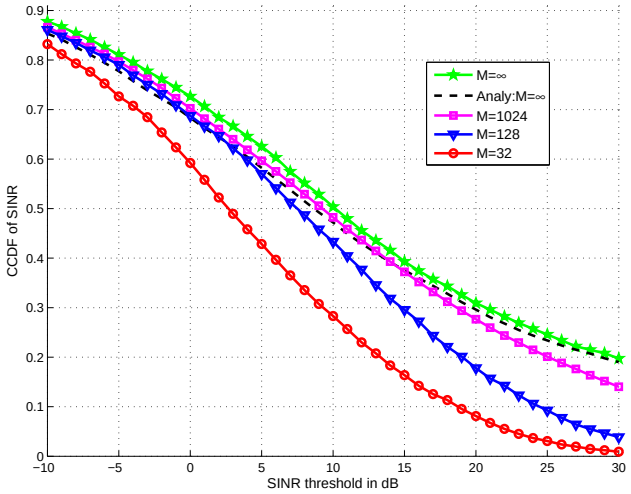
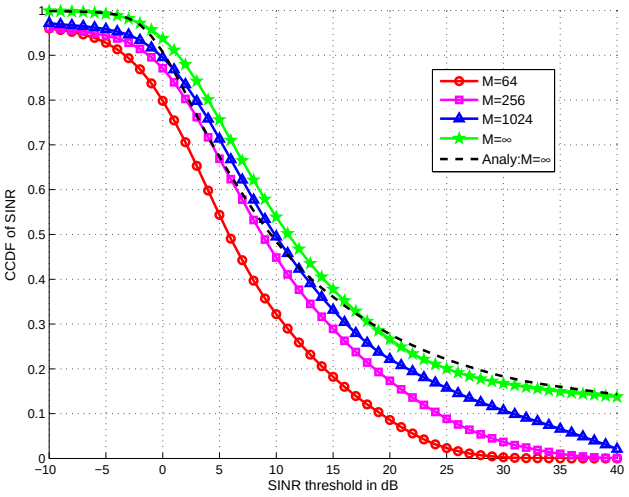


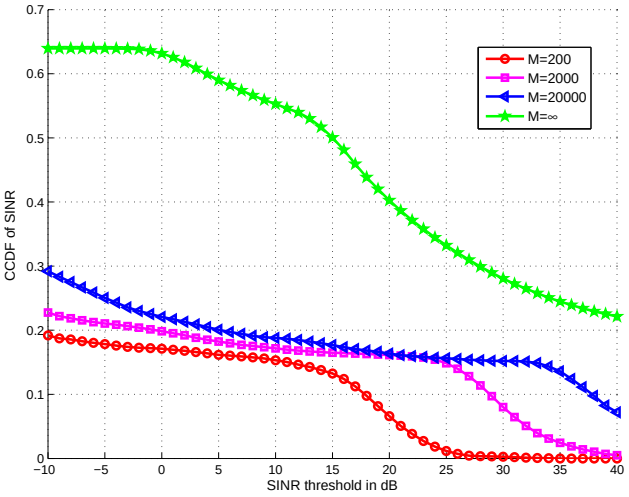
Figure 4.1: Uplink SINR distributions in dense mmWave networks. In the simulations, I assume ISD = 100 meters. The analytical curve is drawn based on Corollary 4.4.1.1.

Fig. 4.1(a), in a dense mmWave network with ISD = 100 meters, the asymptotic SINR distribution provides a close characterization of the SINR curve with $M = 1024$ antennas. Besides, the simulation shows that the analytical expression in Corollary 4.4.1.1 provides a tight approximation.

Next, I present the numerical results for the downlink SINR. In Fig. 4.2, a comparison between the performance in dense and sparse mmWave networks shows that due to the presence of blockages and high noise power, the SINR performance is much sensitive to the base station density, and mmWave massive MIMO networks require dense base station deployments to achieve good coverage. Moreover, the downlink SINR in dense mmWave networks is shown to converge much faster than that in the sparse network. In addition, in Fig. 4.2 (a), I show that Corollary 4.4.2.1 provides a tight approximation



(a) Downlink SINR in dense networks.



(b) Downlink SINR in sparse networks.

Figure 4.2: Downlink mmWave SINR distributions with different base station densities. I assume ISD = 100 m in (a), and ISD = 400 m in (b). The analytical curve in (a) is drawn based on Corollary 4.4.2.1.

of the asymptotic SINR distribution in dense mmWave networks.

Last, I compare the average rate of massive MIMO systems at 2 GHz

Table 4.1: Comparison of Achievable Rates

Carrier	2 GHz	28 GHz	28 GHz
Avg. ISD (m)	500	100	400
Training overhead	20%	14%	14%
Bandwidth (MHz)	100	500	500
Rate per user (Mbps)	52.8	1791.0	436.5
Users per cell	14	4	4
Cell throughput (Mbps)	740.0	7164.0	1745.8

and 28 GHz in Table 4.1. In the rate comparison, I assume the 2 GHz base stations have $M = 64$ antennas, and $W = 100$ MHz bandwidth, while in mmWave systems $M = 128$, and $W = 500$ MHz. Both the uplink and downlink are assumed to take up 50% transmission time. The OFDM technique is assumed to be used in both systems: the parameters for the 2 GHz system are taken from the LTE standard [1], and those for mmWave systems are from [111]. Besides, the training overhead is computed based on [90]. Further details for the rate comparison can be found in [33]. The results show that though serving fewer users per cell, mmWave massive MIMO still outperforms conventional massive MIMO systems in cell throughput, due to the larger bandwidth.

4.6 Conclusion

In this chapter, I analyzed the asymptotic SINR distribution in mmWave massive MIMO networks by incorporated key features of mmWave systems, including the blockage effects and directional beamforming at mobile stations,

into the analytical framework. I provided the asymptotic equivalences for both the uplink and downlink SINR in a large-scale network with Poisson distributed base stations, and derived approximation expressions to compute their distributions. The accuracy of the analytical expressions were verified by numerical simulations. The numerical results showed that mmWave massive MIMO requires a high base station density to achieve good SINR coverage. Moreover, the comparison with massive MIMO systems at lower frequencies showed the promising gain of mmWave massive MIMO over conventional massive MIMO in cell throughput. For future work, it would be interesting to incorporate mmWave hardware constraints, such as hybrid beamforming and one-bit A/D converter, as discussed in Section 1.5.1.

4.7 Appendix

4.7.1 Proof of Corollary 4.4.1

By [26, Lemma 7], conditioning on $\beta_{\ell 0}^{(1)} > 0$, the non-zero path loss process $\left\{1/\beta_{\ell 0}^{(1)}\right\}_{\ell>0}$ is a non-homogenous PPP on $\mathcal{R}^+$ with intensity measure $\Xi(t)$. Moreover, noting that $\beta_{00}^{(1)} = 0$ if and only if the typical user has no LOS nor NLOS base stations, it follows that

$$p_{\mathrm{I}} = \mathbb{P}\left(\beta_{00}^{(1)} > 0\right) = 1 - \mathrm{e}^{-2\pi\lambda_{\mathrm{b}}\int_0^\infty (p_{\mathrm{L}}(x)+p_{\mathrm{N}}(x))x\mathrm{d}x}. \tag{4.7}$$

Therefore, for $T > 0$, the distribution of $1/\beta_{00}^{(1)}$, i.e., the smallest non-zero path loss to the origin, is

$$\mathbb{P}\left(\beta_{00}^{(1)} > T\right) = \mathbb{P}\left(\beta_{00}^{(1)} > T | \beta_{00}^{(1)} > 0\right)\mathbb{P}\left(\beta_{00}^{(1)} > 0\right)$$

$$= \mathbb{P}\left(1/\beta_{00}^{(1)} < 1/T | \beta_{00}^{(1)} > 0\right) \mathbb{P}\left(\beta_{00}^{(1)} > 0\right)$$
$$= \left(1 - \mathrm{e}^{-\Xi(1/T)}\right) p_{\mathrm{I}},$$

where the last step follows from computing the null probability of the PPP formed by the non-zero path loss process $\left\{1/\beta_{\ell 0}^{(1)} > 0\right\}_{\ell > 0}$.

4.7.2 Proof Sketch of Corollary 4.4.1.1

First, for $x > 0$, conditioning on $1/\beta_{00}^{(1)} = x$, I compute the asymptotic uplink SINR distribution as

$$\begin{aligned}
\mathbb{P}\left(\mathrm{SINR}_{\mathrm{UL}} > T | 1/\beta_{00}^{(1)} = x\right) &= \mathbb{P}\left(1 > \frac{Tx^2}{Q^2}\left(\sum_{\ell \neq 0} D_{0\ell}^{(1)2} \beta_{0\ell}^{(1)2}\right)\right) \\
&\overset{(a)}{\approx} \mathbb{P}\left[\gamma > \frac{Tx^2}{Q^2}\left(\sum_{\ell \neq 0} D_{0\ell}^{(1)2} \beta_{0\ell}^{(1)2}\right)\right] \\
&\overset{(b)}{\approx} 1 - \mathbb{E}_{\mathcal{N}_{\mathrm{u}}^{(1)}}\left[\left(1 - \mathrm{e}^{-\eta \frac{Tx^2}{Q^2}\left(\sum_{\ell \neq 0} D_{0\ell}^{(1)2} \beta_{0\ell}^{(1)2}\right)}\right)^N\right] \\
&= \sum_{n=1}^{N} \binom{N}{n} (-1)^{n+1} \mathbb{E}_{\mathcal{N}_{\mathrm{u}}^{(1)}}\left[\mathrm{e}^{-n\eta \frac{Tx^2}{Q^2}\left(\sum_{\ell \neq 0} D_{0\ell}^{(1)2} \beta_{0\ell}^{(1)2}\right)}\right] \\
&\overset{(b)}{=} \sum_{n=1}^{N} \binom{N}{n} (-1)^{n+1} \mathrm{e}^{-W_n(T,x) - V_n(T,x)}
\end{aligned}$$

where in (a) γ is a Gamma random variable with unit mean and shape parameter N, and the approximation follows from the fact $\lim_{n\to\infty} \frac{n^n x^{n-1} \mathrm{e}^{-nx}}{\Gamma(n)} = \delta(x-1)$ [20]; in (b), $\eta = N(N!)^{1/N}$, and the step follows from Alzer's inequality [14]; (c) follows from computing the Laplace functional of the inhomogeneous PPP $\mathcal{N}_{\mathrm{u}}^{(1)}$, whose intensity function is given in Approximation 4.3.1. Last,

the proof is completed by deconditioning on $1/\beta_{00}^{(1)} = x$ for $x > 0$, whose distribution is given in Lemma 4.4.1.

Chapter 5

Conclusion

5.1 Summary

In this dissertation, I developed mathematical frameworks for mmWave, sub-6 GHz massive MIMO, and mmWave massive MIMO cellular networks, using stochastic geometry. The proposed analytical results reveal quantitative relationships between key system parameters in 5G cellular networks. In my first contribution, I proposed a general system model for mmWave cellular networks that incorporates blockage effect and directional beamforming, and analyzed the downlink SINR and rate performance. In my second contribution, I analyzed the uplink SINR and rate as a function of the number of antennas and scheduled users per cell in massive MIMO networks. In my third contribution, I extend the massive MIMO network model to mmWave frequencies, and analyzed the asymptotic SINR and rate distributions, which are representative of the performance with finite large antenna arrays in certain massive MIMO networks. The contributions presented in this dissertation demonstrate key performance trends, and provide important system design insights in 5G cellular networks.

To evaluate the coverage and rate performance in mmWave cellular

networks, I proposed a general framework that incorporates blockage effect and directional beamforming in a tractable manner. I applied a distance-dependent line-of-site (LOS) probability function, and divided the locations of the LOS and non-LOS base stations into two independent non-homogeneous Poisson point processes, to which different path loss laws are applied. Based on the proposed framework, I examined the mmWave coverage and rate distributions as a function of the antenna geometry and base station density. I further investigated the case of dense networks by applying a simplified system model, in which the LOS region of a user is approximated as a fixed LOS ball. I showed that dense mmWave networks can achieve comparable coverage and much higher data rates than sub-6 GHz systems, despite the presence of blockages. I also prove that increasing base station density beyond a critical point will harm the SINR and rate performance, and that the cell size to achieve best optimal SINR scales with the average size of the area unblocked by blockages to a user.

To analyze the SINR and rate performance in a large-scale uplink massive MIMO network, I developed a stochastic geometry framework that incorporates the impacts of fractional power control and pilot contamination. Based on the model, I derived expressions for the spatial average SINR distributions over user and base station distributions with MRC and ZF receivers. I show that using massive MIMO, the uplink SINR in certain urban marco-cell scenarios is limited by interference. In the interference-limited regime, the results reveal that for MRC receivers, a super-linear (polynomial) scal-

ing law between the number of base station antennas and scheduled users per cell preserves the uplink SIR distribution, while a linear scaling applies to ZF receivers. Based on the scaling results, I show that ZF receivers outperform MRC receivers in the SIR coverage, and quantified the performance gap in terms of the difference in the number of antennas to achieve the same SIR distribution. In addition, I demonstrate that the optimal compensation fraction in fractional power control to optimize rate is generally different for MRC and ZF receivers. Besides, I show in simulations that the scaling results derived from the proposed framework apply to the networks where base stations are distributed according to a hexagonal lattice.

To compare the performance between sub-6 GHz and mmWave massive MIMO, I proposed a stochastic geometry framework for mmWave massive MIMO networks, where the differences in propagation and hardware constraints were considered. I focused on the asymptotic regime (in the limit of the number of antennas), and derived analytical expressions for the asymptotic SINR distributions in both uplink and downlink. I used numerical simulations to show a fast convergence to its asymptotic SINR in mmWave massive MIMO networks, when the base station density is sufficiently high. Finally, I compared the cell throughput of mmWave massive MIMO and sub-6 GHz massive MIMO, and concluded that the optimum choice of carrier frequency depends on the base station density.

5.2 Future Research Directions

In this section, I present some promising directions for future work related to my contributions in this dissertation.

1. **Uplink SINR and rate analysis in mmWave cellular networks**

   The contribution in Chapter 2 focus on the downlink SINR performance. It is natural to consider to analyze the uplink SINR using a similar framework. Prior analysis in sub-6 GHz cellular networks shows different topologies of interferers between downlink and uplink [108, 130]. For example, when the interfering base stations are assumed to be distributed as a PPP, then the scheduled users, i.e. the interferers in the uplink, in all cells do not form a PPP, due to the Voronoi cell structure [108, 130]. Therefoere, it is essential to incorporate such difference in the uplink analysis in mmWave cellular networks. One promising approach to compute the uplink SIR distribution is to extend the approach in [130] that characterizes the pairwise correlations in the uplink interferers to mmWave networks.

2. **3D system model for mmWave networks**

   In reality, the heights of blockages, base stations antennas, or even mobile users may have non-negligible impacts on system performance. Thus, it is of interest to incorporate the elevation height in analysis and extend the proposed framework in Chapter 2 into 3D space. The 3D system

model will also make it possible to analyze the performance of the techniques that exploit the spatial degrees of freedom in elevation directions, such as 3D beamforming and vertical secterization [2].

3. **Performance analysis for indoor users**

   Prior analyses in Chapter 2 mainly focus on the coverage of outdoor users. Due to the large penetration loss, outdoor mmWave base stations will hardly serve an indoor user. The indoor users can be served by mmWave small cells (distributed antennas) inside buildings, or by base stations operating at lower carrier frequencies. It would be interesting to examine the performance of mmWave systems with indoor infrastructures co-located with buildings.

4. **Modeling and analysis of multi-band system**

   To overcome the mmWave coverage holes due to blockages, one promising design is to have a joint sub-6 GHz-mmWave system, where the mmWave base stations act as hot spots to provide high data rates for the users in the range, and the sub-6 GHz macro base stations provide a robust coverage. Even though, the interference powers at two frequencies are independent, the SINR distributions can be correlated in reality. For one thing, some of the mmWave and sub-6 GHz base stations may be co-located. Consequently, the gain of sub-6 GHz-mmWave joint coverage can be overestimated by assuming independent BS locations. For another, the load of mmWave and sub-6 GHz can be correlated, when

serving the same group of users. Thus, if the load is balanced between two systems, the number of active interfering nodes in each tier are also correlated, which breaks the independence of the SINR coverage between tiers. In future work, novel modeling approaches that take account for the correlations in both base stations locations and load are essential to understand the gain of such multi-band joint-coverage system.

5. **Scaling law for downlink massive MIMO SIR**

   Only the uplink SIR was analyzed based on the system model in Chapter 3. The approach does not directly apply to analyze the downlink performance, as the precoders in the downlink are normalized to address the power constraint, while the normalization of combiner vector does not change the uplink SIR distribution. The normalization factor makes the downlink SIR distribution intractable to compute, as it brings in (high) correlations between certain terms in the SIR expression. One potential solution is to neglect the correlations between certain out-of-cell interference terms, derive an approximate downlink SIR distribution that could reasonable match the simulations, and derive the scaling law based on the approximation.

6. **System analysis of sparsity-based algorithms in mmWave networks**

   Due to the channel sparsity, mmWave systems will unavoidably apply compressed-sensing based signal processing techniques to estimate chan-

nel and design precoder/ combiner [122]. The performance of the sparsity algorithms, e.g. in [12, 93], have not yet been examined on a network level. One promising approach is to formulate the signal and interference power as random functions of the system parameter, e.g. the base station and blockage densities, and then apply the signal and interference distributions to compute the system-level performance for the sparsity-based algorithms. For example, treating the interference as noise in channel estimations, then it is possible to derive the equivalent SNR distribution in the sparsity-recovery problem in [93], and compute the channel estimation error at a typical user, when taking accounts for the random distributions of the serving and interfering base stations.

Bibliography

[1] 3GPP TR 36.814. Further advancements for E-UTRA physical layer aspects (Release 9). Mar. 2010.

[2] 3GPP TR 36.873. Technical Specification Group Radio Access Network; Study on 3D Channel MOdel for LTE (Release 12). Sep. 2014.

[3] 3GPP TR 36.942. Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios (Release 11). Mar. 2012.

[4] A. Adhikary, Junyoung Nam, Jae-Young Ahn, and G. Caire. Joint spatial division and multiplexing: The large-scale array regime. *IEEE Trans. on Information Theory*, 59(10):6441–6463, Oct. 2013.

[5] M. R. Akdeniz et al. Millimeter wave picocellular system evaluation for urban deployments. *arXiv preprint arXiv:1304.3963*, 2013.

[6] M.R. Akdeniz, Yuanpeng Liu, M.K. Samimi, Shu Sun, S. Rangan, T.S. Rappaport, and E. Erkip. Millimeter wave channel modeling and cellular capacity evaluation. *IEEE J. Sel. Areas Commun.*, 32(6):1164–1179, June 2014.

[7] S. Akoum, E. O. Ayach, and R. W. Heath Jr. Coverage and capacity in mmWave cellular systems. In *Proc. of the Forty Sixth Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, pages 688–692, 2012.

[8] A.V. Alejos et al. Measurement and analysis of propagation mechanisms at 40 GHz: Viability of site shielding forced by obstacles. *IEEE Trans. Veh. Technol.*, 57(6):3369–3380, 2008.

[9] A. Alkhateeb, O. El Ayach, G. Leus, and R.W. Heath Jr. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):831–846, Oct. 2014.

[10] A. Alkhateeb, Jianhua Mo, N. Gonzalez-Prelcic, and R.W. Heath. Mimo precoding and combining solutions for millimeter-wave systems. *IEEE Communications Magazine*, 52(12):122–131, Dec. 2014.

[11] A. Alkhateeb, Jianhua Mo, N. Gonzalez-Prelcic, and R.W. Heath. MIMO precoding and combining solutions for millimeter-wave systems. *IEEE Communications Magazine*, 52(12):122–131, Dec. 2014.

[12] A. Alkhateeb, O., G. Leus, and R.W. Heath Jr. Compressed sensing based multi-user millimeter wave systems: how many measurements are needed? In *Proc. of IEEE Int. Conf. on Commun. (ICC)*, June 2015.

[13] Ahmed Alkhateeb and Robert W Heath Jr. Frequency selective hybrid precoding for limited feedback millimeter wave systems. *submitted to*

IEEE Transactions on Communications, arXiv preprint arXiv:1510.00609, 2015.

[14] Horst Alzer. On some inequalities for the incomplete Gamma function. *Mathematics of Computation*, 66(218):771–778, 1997.

[15] C.R. Anderson and T.S. Rappaport. In-building wideband partition loss measurements at 2.5 and 60 GHz. *IEEE Trans. Wireless Commun.*, 3(3):922–928, 2004.

[16] Jeffrey G. Andrews, Francois Baccelli, and Radha Krishna Ganti. A tractable approach to coverage and rate in cellular networks. *IEEE Trans. Commun.*, 59(11):3122–3134, Nov. 2011.

[17] J.G. Andrews, S. Buzzi, Wan Choi, S.V. Hanly, A. Lozano, A.C.K. Soong, and J.C. Zhang. What will 5G be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, June 2014.

[18] J.G. Andrews et al. A tractable approach to coverage and rate in cellular networks. *IEEE Trans. Commun.*, 59(11):3122–3134, 2011.

[19] K. Appaiah, A. Ashikhmin, and T.L. Marzetta. Pilot contamination reduction in multi-user tdd systems. In *Proc. of ICC*, pages 1–5, May 2010.

[20] R. Aris. *Mathematical Modeling: A Chemical Engineer's Perspective*. Academic Press, 1999.

[21] A. Ashikhmin and T. Marzetta. Pilot contamination precoding in multi-cell large scale antenna systems. In *Proc. of IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 1137–1141, 2012.

[22] I. Atzeni, J. Arnau, and M. Debbah. Fractional pilot reuse in massive mimo systems. In *IProc. of IEEE ICC Workshops*, pages 1030–1035, June 2015.

[23] Omar El Ayach, Sridhar Rajagopal, Shadi Abu-Surra, Zhouyue Pi, and Robert W Heath Jr. Spatially sparse precoding in millimeter wave MIMO systems. *To appear in IEEE Trans. Wireless Commun.*, 2013.

[24] F. Baccelli and B. Blaszczyszyn. *Stochastic geometry and wireless networks. Volume I: theory.* NOW publishers, 2009.

[25] T. Bai, A. Alkhateeb, and R. W. Heath Jr. Coverage and capacity in millimeter wave cellular networks. *IEEE Commun. Mag.*, Sep. 2014.

[26] T. Bai and R. W. Heath Jr. Coverage analysis for millimeter wave cellular networks with blockage effects. In *Proc. of IEEE Global Conf. on Signal and Information Processing (GlobalSIP)*, Dec. 2013.

[27] T. Bai and R. W. Heath Jr. Asymptotic coverage probability and rate in massive MIMO networks. In *Proc. of IEEE Global Conf. on Signal and Information Processing (GlobalSIP)*, Dec. 2014.

[28] T. Bai and R. W. Heath Jr. Coverage and rate analysis for millimeter wave cellular networks. March 2014.

[29] T. Bai and R.W. Heath Jr. Coverage and rate analysis for millimeter-wave cellular networks. *IEEE Trans. Wireless Commun.*, 14(2):1100–1114, Feb. 2015.

[30] T. Bai, R. Vaze, and R. W. Heath Jr. Analysis of blockage effects on urban cellular networks. *IEEE Trans.Wireless Commun.*, 13(9):5070–5083, Sep. 2014.

[31] Tianyang Bai and R. W. Heath Jr. Coverage in dense millimeter wave cellular networks. In *Proc. of the Forty Seventh Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, pages 1–5, Nov. 2013.

[32] Tianyang Bai and R. W. Heath Jr. Analysis of self-body blocking effects in millimeter wave cellular networks. In *Proc. of Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, Nov. 2014.

[33] Tianyang Bai and Robert W. Heath Jr. Massive MIMO: Millimeter wave or lower frequencies? *Preprint*, Feb. 2015.

[34] Tianyang Bai and Robert W Heath, Jr. Uplink massive MIMO SIR analysis: How do antennas scale with users? In *Proc. of IEEE Globecom Conf.*, 2015.

[35] T. Baykas et al. IEEE 802.15.3c: the first IEEE wireless standard for data rates over 1 Gb/s. *IEEE Commun. Mag.*, 49(7):114–121, 2011.

[36] D. E. Berraki, S. M. D. Armour, and A. R. Nix. Application of compressive sensing in sparse spatial channel recovery for beamforming in

mmWave outdoor systems. In *Proc. of WCNC*, pages 887–892, April 2014.

[37] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah. Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits. *IEEE Transactions on Information Theory,*, 60(11):7112–7139, Nov. 2014.

[38] E. Bjornson, M. Matthaiou, and M. Debbah. Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design. *IEEE Trans. on Wireless Commun.*, 14(8):4353–4368, Aug. 2015.

[39] Emil Björnson, Erik G. Larsson, and Mérouane Debbah. Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated? 2014.

[40] Emil Björnson, Luca Sanguinetti, and Marios Kountouris. Deploying dense networks for maximal energy efficiency: Small cells meet massive MIMO. 2015.

[41] B. Blaszczyszyn, M. K. Karray, and H. P. Keeler. Using Poisson processes to model lattice cellular networks. In *Prof. of IEEE International Conference on Computer Communications (INFOCOM)*, 2013.

[42] F Boccardi, R.W. Heath, A. Lozano, T.L. Marzetta, and P. Popovski. Five disruptive technology directions for 5G. *IEEE Communications Magazine*, 52(2):74–80, February 2014.

[43] J. Brady, N. Behdad, and A. M. Sayeed. Beamspace mimo for millimeter-wave communications: System architecture, modeling, analysis, and measurements. *IEEE Transactions on Antennas and Propagation*, 61(7):3814–3827, July 2013.

[44] G. Caire, Sean A. Ramprashad, and H.C. Papadopoulos. Rethinking network mimo: Cost of csit, performance analysis, and architecture comparisons. In *Proc. of Information Theory and Applications Workshop (ITA)*, pages 1–10, Jan. 2010.

[45] V. Chandrasekhar, M. Kountouris, and J.G. Andrews. Coverage in multi-antenna two-tier networks. *IEEE Transactions on Wireless Commun.*, 8(10):5314–5327, Oct. 2009.

[46] Junil Choi, Jianhua Mo, and Robert W. Heath Jr. Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit adcs. 2015.

[47] Federal Communication Commission. *FCC 15-138 NOTICE OF PROPOSED RULEMAKING*. 2015.

[48] The WirelessHD Consortium. *WirelessHD Specification Overview*. 2010.

[49] Deliverable D5.3 V1.0. WINNER+ final channel models. Mar. 2010.

[50] H. S. Dhillon et al. Modeling and analysis of K-tier downlink heterogeneous cellular networks. *IEEE Journal on Selected Areas in Communications*, 30(3):550–560, April 2012.

[51] H.S. Dhillon, M. Kountouris, and J.G. Andrews. Downlink MIMO HetNets: Modeling, ordering results and performance analysis. *IEEE Trans. Wireless Commun.*, 12(10):5208–5222, October 2013.

[52] O. El Ayach, S. Rajagopal, S. Abu-Surra, Zhouyue Pi, and R.W. Heath. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 13(3):1499–1513, Mar. 2014.

[53] H. ElSawy and E. Hossain. On stochastic geometry modeling of cellular uplink transmission with truncated channel inversion power control. *IEEE Trans. Wireless Commun.*, 13(8):4454–4469, Aug. 2014.

[54] D.T. Emerson. The work of Jagadis Chandra Bose: 100 years of millimeter-wave research. *IEEE Transactions on Microwave Theory and Techniques,*, 45(12):2267–2273, Dec. 1997.

[55] M. Franceschetti, J. Bruck, and L.J. Schulman. A random walk model of wave propagation. *IEEE Trans. Antennas Propag.*, 52(5):1304–1317, 2004.

[56] Xiang Gao, O. Edfors, F. Rusek, and F. Tufvesson. Linear pre-coding performance in measured very-large mimo channels. In *Proc. of IEEE VTC Fall*, pages 1–5, Sept. 2011.

[57] Xiang Gao, O. Edfors, F. Rusek, and F. Tufvesson. Massive mimo performance evaluation based on measured propagation data. *IEEE Trans. Wireless Commun.*, 14(7):3899–3911, July 2015.

[58] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas. LTE-advanced: next-generation wireless broadband technology. *IEEE Wireless Communications*, 17(3):10–22, June 2010.

[59] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[60] L.J. Greenstein and V. Erceg. Gain reductions due to scatter on wireless paths with directional antennas. *IEEE Commun. Let.*, 3(6):169–171, 1999.

[61] Shuangfeng Han, Chih-Lin I, Zhikun Xu, and C. Rowell. Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G. *IEEE Communications Magazine*, 53(1):186–194, Jan. 2015.

[62] R. W. Heath Jr. Coverage and capacity analysis of mmWave cellular systems. *Presentation delivered at Int. Conf. on Communi. (ICC)*, 2013.

[63] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink. Channel measurements for large antenna arrays. In *Proc. of Interntation Symposium on Wireless Communication Systems*, pages 811–815, Aug. 2012.

[64] J. Hoydis, S. ten Brink, and M. Debbah. Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? *IEEE J. Sel. Areas Commun.*, 31(2):160–171, Feb. 2013.

[65] Hoon Huh, G. Caire, H.C. Papadopoulos, and S.A. Ramprashad. Achieving "Massive MIMO" spectral efficiency with a not-so-large number of antennas. *IEEE Trans. on Wireless Commun.*, 11(9):3226–3239, September 2012.

[66] A.M. Hunter, J.G. Andrews, and S. Weber. Transmission capacity of ad hoc networks with spatial diversity. *IEEE Trans. Wireless Commun.*, 7(12):5058–5071, December 2008.

[67] Sooyoung Hur, Taejoon Kim, D.J. Love, J.V. Krogmeier, T.A. Thomas, and A. Ghosh. Millimeter wave beamforming for wireless backhaul and access in small cell networks. *IEEE Trans. Commun.*, 61(10):4391–4403, Oct. 2013.

[68] IEEE. IEEE Standard - Part 11: Wireless LAN MAC and PHY Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band. 2012.

[69] IEEE. IEEE Standard for WirelessMAN-Advanced Air Interface for Broadband Wireless Access Systems. 2012.

[70] J. Jose, A. Ashikhmin, T.L. Marzetta, and S. Vishwanath. Pilot contamination and precoding in multi-cell TDD systems. *IEEE Trans. Wireless Commun.*, 10(8):2640–2651, Aug. 2011.

[71] M. Kiessling and J. Speidel. Analytical performance of MIMO zero-forcing receivers in correlated Rayleigh fading environments. In *Proc.*

of SPAWC, pages 383–387, June 2003.

[72] M. Kountouris and J.G. Andrews. Downlink SDMA with limited feedback in interference-limited wireless networks. *IEEE Trans. Wireless Commun.*, 11(8):2730–2741, Aug. 2012.

[73] N. Krishnan, R.D. Yates, and N.B. Mandayam. Uplink linear receivers for multi-cell multiuser MIMO with pilot contamination: Large system analysis. *IEEE Trans. Wireless Commun.*, 13(8):4360–4373, Aug. 2014.

[74] S. Lakshminarayana, M. Assaad, and M. Debbah. Coordinated multicell beamforming for massive mimo: A random matrix approach. *IEEE Transactions on Information Theory*, 61(6):3387–3412, June 2015.

[75] B. Langen, G. Lober, and W. Herzig. Reflection and transmission behaviour of building materials at 60 GHz. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 2, pages 505–509, Sep. 1994.

[76] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, Feb. 2014.

[77] Chang Li, Jun Zhang, and K.B. Letaief. Performance analysis of SDMA in multicell wireless networks. In *In Proc of Globecom*, pages 3867–3872, Dec. 2013.

[78] Liangbin Li, A. Ashikhmin, and T. Marzetta. Pilot contamination precoding for interference reduction in large scale antenna systems. In *Proc. of 51st Annual Allerton Conference on Communication, Control, and Computing*, pages 226–232, Oct. 2013.

[79] Xueru Li, Shidong Zhou, E. Bjornson, and Jing Wang. Capacity analysis for spatially non-wide sense stationary uplink massive MIMO systems. *IEEE Trans. on Wireless Commun.*, 14(12):7044–7056, Dec. 2015.

[80] Zheda Li, Rui Wang, and A.F. Molisch. Shadowing in urban environments with microcellular or peer-to-peer links. In *Proc. of 6th European Conf. Antennas and Propag. (EUCAP)*, pages 44–48, 2012.

[81] Le Liang, Wei Xu, and Xiaodai Dong. Low-complexity hybrid precoding in massive multiuser MIMO systems. *IEEE Wireless Communications Letters*, 3(6):653–656, Dec. 2014.

[82] Ning Liang, Wenyi Zhang, and Cong Shen. An uplink interference analysis for massive MIMO systems with MRC and ZF receivers. 2015.

[83] Yeon-Geun Lim, Chan-Byoung Chae, and G. Caire. Performance analysis of massive mimo for cell-boundary users. *IEEE Trans. Wireless Commun.*, 14(12):6827–6842, Dec. 2015.

[84] Lingfeng Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P.D. Doncker. The cost 2100 mimo channel model. *IEEE Wireless Communications*, 19(6):92–99, Dec. 2012.

[85] S.L. Loyka. Channel capacity of MIMO architecture using the exponential correlation matrix. *IEEE Commun. Let.*, 5(9):369–371, Sep. 2001.

[86] Jonathan S. Lu, Daniel Steinbach, Patrick Cabrol, and Philip Pietraski. Modeling human blockers in millimeter wave radio links. *ZTE Communications*, 10(4):23–28, Dec. 2012.

[87] Lu Lu, G.Y. Li, A.L. Swindlehurst, A. Ashikhmin, and Rui Zhang. An overview of massive MIMO: Benefits and challenges. *IEEE J. Sel. Topics Signal Process.*, 8(5):742–758, Oct. 2014.

[88] U. Madhow, D.R. Brown, S. Dasgupta, and R. Mudumbai. Distributed massive mimo: Algorithms, architectures and concept systems. In *Proc. of Information Theory and Applications Workshop (ITA)*, pages 1–7, Feb 2014.

[89] P. Madhusudhanan, Xing Li, Youjian Liu, and T.X. Brown. Stochastic geometric modeling and interference analysis for massive MIMO systems. In *Proc.of Modeling Optimization in Mobile, Ad Hoc Wireless Networks (WiOpt),*, pages 15–22, May 2013.

[90] T.L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans. Wireless Commun.*, 9(11):3590–3600, Nov. 2010.

[91] C. Masouros and M. Matthaiou. Space-constrained massive MIMO: Hitting the wall of favorable propagation. *IEEE Communications Letters*, 19(5):771–774, May 2015.

[92] H. Mehrpouyan, M.R. Khanzadi, M. Matthaiou, A.M. Sayeed, R. Schober, and Yingbo Hua. Improving bandwidth efficiency in E-band communication systems. *IEEE Communications Magazine*, 52(3):121–128, March 2014.

[93] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R.W. Heath. Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches? *IEEE Access*, PP(99):1–1, 2016.

[94] A. Mezghani and J. A. Nossek. On ultra-wideband mimo systems with 1-bit quantized outputs: Performance analysis and input optimization. In *Proc. of ISIT*, pages 1286–1289, June 2007.

[95] J. Mo and R. W. Heath. Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information. *IEEE Transactions on Signal Processing*, 63(20):5498–5512, Oct. 2015.

[96] J.. Mo and R. W. Heath Jr. High SNR capacity of millimeter wave MIMO systems with one-bit quantization. In *Proc. of Information Theory and Applications (ITA)*, Feb. 2014.

[97] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath. Channel estimation in millimeter wave mimo systems with one-bit quantization. In *Proc.*

of Asilomar Conference on Signals, Systems and Computers, pages 957–961, Nov 2014.

[98] S.K. Mohammed and E.G. Larsson. Per-antenna constant envelope precoding for large multi-user MIMO systems. *IEEE Trans. on Commun.*, 61(3):1059–1071, March 2013.

[99] Christopher Mollén, Junil Choi, Erik G. Larsson, and Robert W. Heath Jr. One-bit ADCs in wideband massive MIMO systems with OFDM transmission. In *Proc. of IEEE ICASSP*, March 2016.

[100] R.R. Muller, L. Cottatellucci, and M. Vehkapera. Blind pilot decontamination. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):773–786, Oct. 2014.

[101] B. Murmann. ADC performance survey 1997-2015. 2015.

[102] A. Y. Nashashibi, K. Sarabandi, S. Oveisgharan, M. C. Dobson, W. S. Walker, and E. Burke. Millimeter-wave measurements of foliage attenuation and ground reflectivity of tree stands at nadir incidence. *IEEE Transactions on Antennas and Propagation*, 52(5):1211–1222, May 2004.

[103] Hien Quoc Ngo and E.G. Larsson. Evd-based channel estimation in multicell multiuser mimo systems with very large antenna arrays. In *In Proc. of IEEE Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3249–3252, March 2012.

[104] Hien Quoc Ngo, E.G. Larsson, and T.L. Marzetta. Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Trans. on Commun.*, 61(4):1436–1449, April 2013.

[105] Hien Quoc Ngo, E.G. Larsson, and T.L. Marzetta. Aspects of favorable propagation in massive MIMO. In *Proc. of European Signal Processing Conf. (EUSIPCO)*, pages 76–80, Sep. 2014.

[106] Hien Quoc Ngo, M. Matthaiou, T.Q. Duong, and E.G. Larsson. Uplink performance analysis of multicell MU-SIMO systems with ZF receivers. *IEEE Trans. Veh. Tech.*, 62(9):4471–4483, Nov. 2013.

[107] Weiheng Ni and Xiaodai Dong. Hybrid block diagonalization for massive multiuser MIMO systems. *IEEE Trans. on Commun.*, 64(1):201–211, Jan. 2016.

[108] T.D. Novlan, H.S. Dhillon, and J.G. Andrews. Analytical modeling of uplink cellular networks. *IEEE Trans. Wireless Commun.*, 12(6):2669–2679, June 2013.

[109] S. Payami and F. Tufvesson. Channel measurements and analysis for very large array systems at 2.6 GHz. In *Proc. of European Conf. Antennas and Propag.*, pages 433–437, March 2012.

[110] Yiyang Pei, The-Hanh Pham, and Ying-Chang Liang. How many RF chains are optimal for large-scale MIMO systems when circuit power is

considered? In *Proc. of Global Communications Conference (GLOBECOM)*, pages 3868–3873, Dec 2012.

[111] Z Pi and F Khan. A millimeter-wave massive MIMO system for next generation mobile broadband. In *Proc. of the Forty Sixth Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, pages 693–698, 2012.

[112] Zhouyue Pi and F. Khan. An introduction to millimeter-wave mobile broadband systems. *IEEE Communications Magazine*, 49(6):101–107, June 2011.

[113] S. Rajagopal, S. Abu-Surra, and M. Malmirchegini. Channel feasibility for outdoor non-line-of-sight mmwave mobile communication. In *Proc. of IEEE Veh. Technol. Conf. (VTC Fall)*, pages 1–6, 2012.

[114] D. Ramasamy, S. Venkateswaran, and U. Madhow. Compressive tracking with 1000-element arrays: A framework for multi-gbps mm wave cellular downlinks. In *In Proc. of Allerton Conf. on Communication, Control, and Computing*, pages 690–697, Oct. 2012.

[115] Sundeep Rangan. Generalized approximate message passing for estimation with random linear mixing. abs/1010.5141, 2010.

[116] Theodore S. Rappaport, Robert W. Heath, Jr., Robert C. Daniels, and James N. Murdock. *Millimeter Wave Wireless Communications*. Prentice Hall, 2014.

[117] T.S. Rappaport and Sijia Deng. 73 ghz wideband millimeter-wave foliage and ground reflection measurements and models. In *2015 IEEE International Conference on Communications Workshops*, pages 1238–1243, June 2015.

[118] T.S. Rappaport and et al. 38 GHz and 60 GHz angle-dependent propagation for cellular and peer-to-peer wireless communications. In *Proc. of IEEE Int. Conf. on Commun. (ICC)*, pages 4568–4573, Jun. 2012.

[119] T.S. Rappaport et al. Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications. *IEEE Trans. on Antennas Propag.*, 61(4):1850–1859, Apr. 2013.

[120] T.S. Rappaport et al. Millimeter wave mobile communications for 5G cellular: It will work! *IEEE Access*, 1:335–349, 2013.

[121] T.S. Rappaport, G.R. Maccartney, M.K. Samimi, and Shu Sun. Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE Transactions on Communications*, 63(9):3029–3056, Sept. 2015.

[122] Robert W. Heath Jr., Nuria González Prelcic, Sundeep Rangan, Wonil Roh, and Akbar M. Sayeed. An overview of signal processing techniques for millimeter wave MIMO systems. 2015.

[123] R. Rogalin, O.Y. Bursalioglu, H. Papadopoulos, G. Caire, A.F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis. Scalable synchronization and reciprocity calibration for distributed multiuser MIMO. *IEEE Transactions on Wireless Communication*, 13(4):1815–1831, April 2014.

[124] F. Rusek, D. Persson, Buon Kiong Lau, E.G. Larsson, T.L. Marzetta, O. Edfors, and F. Tufvesson. Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1):40–60, Jan 2013.

[125] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno. Coordinated multipoint transmission/reception techniques for LTE-advanced. *IEEE Wireless Communications*, 17(3):26–34, June 2010.

[126] S.Y. Seidel and T.S. Rappaport. Site-specific propagation prediction for wireless in-building personal communication system design. *IEEE Trans. Veh. Technol.*, 43(4):879–891, 1994.

[127] Clayton Shepard, Hang Yu, Narendra Anand, Erran Li, Thomas Marzetta, Richard Yang, and Lin Zhong. Argos: Practical many-antenna base stations. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, Mobicom '12, pages 53–64, New York, NY, USA, 2012. ACM.

[128] J. Singh, S. Ponnuru, and U. Madhow. Multi-Gigabit communication: the ADC bottleneck. In *IEEE International Conference on Ultra-Wideband (ICUWB).*, pages 22–27, Sep. 2009.

[129] S. Singh, H.S. Dhillon, and J.G. Andrews. Offloading in heterogeneous networks: Modeling, analysis, and design insights. *Wireless Communications, IEEE Transactions on*, 12(5):2484–2497, May 2013.

[130] S. Singh, X. Zhang, and J. Andrews. Joint rate and SINR coverage analysis for decoupled uplink-downlink biased cell associations in HetNets. *To appear in IEEE Trans. on Wireless Commun.*, 2015.

[131] S. Singh, F. Ziliotto, U. Madhow, E. Belding, and M. Rodwell. Blockage and directivity in 60 GHz wireless personal area networks: from cross-layer model to multihop MAC design. *IEEE Journal on Selected Areas in Communications*, 27(8):1400–1413, October 2009.

[132] C. Studer and E.G. Larsson. Par-aware large-scale multi-user MIMO-OFDM downlink. *IEEE Journal on Selected Areas in Communications*, 31(2):303–313, February 2013.

[133] A.L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino. Millimeter-wave massive MIMO: the next wireless revolution? *Communications Magazine, IEEE*, 52(9):56–62, September 2014.

[134] C. Tepedelenlioglu, A. Rajan, and Yuan Zhang. Applications of stochastic ordering to wireless communications. *IEEE Trans. Wireless Commun.*, 10(12):4249–4257, December 2011.

[135] A. Toscano, F. Bilotti, and L. Vegni. Fast ray-tracing technique for electromagnetic field prediction in mobile communications. *IEEE Trans.*

on Magnetics, 39(3):1238–1241, May 2003.

[136] K.T. Truong and R.W. Heath. Effects of channel aging in massive MIMO systems. *Journal of Communications and Networks*, 15(4):338–351, Aug. 2013.

[137] K.T. Truong and R.W. Heath Jr. The viability of distributed antennas for massive mimo systems. In *Proc. of Asilomar Conf. on Signals, Systems and Computers*, pages 1318–1323, Nov. 2013.

[138] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[139] J. Vieira, F. Rusek, and F. Tufvesson. Reciprocity calibration methods for massive mimo based on antenna coupling. In *Proc. of IEEE Global Communications Conference*, pages 3708–3712, Dec. 2014.

[140] R. H. Walden. Analog-to-digital converter survey and analysis. *IEEE Journal on Selected Areas in Communications*, 17(4):539–550, April 1999.

[141] J. Wang, Z. Lan, C.W. Pyo, T. Baykas, C.S. Sum, M.A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, et al. Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems. *IEEE Journal on Selected Areas in Communications*, 27(8):1390–1399, Nov. 2009.

[142] Shengchu Wang, Yunzhou Li, and Jing Wang. Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs. *IEEE Trans. Wireless Communn.*, 14(4):2156–2168, April 2015.

[143] Shangbin Wu, Cheng-Xiang Wang, E.-H.M. Aggoune, M.M. Alwakeel, and Yejun He. A non-stationary 3D wideband twin-cluster model for 5G massive MIMO channels. *IEEE Journal on Selected Areas in Communications*, 32(6):1207–1218, June 2014.

[144] Shangbin Wu, Cheng-Xiang Wang, H. Haas, E.-H.M. Aggoune, M.M. Alwakeel, and Bo Ai. A non-stationary wideband channel model for massive MIMO communication systems. *IEEE Trans. on Wireless Commun.*, 14(3):1434–1446, March 2015.

[145] Y. Wu, Y. Cui, and B. Clerckx. Analysis and optimization of inter-tier interference coordination in downlink multi-antenna HetNets with offloading. *To appear in IEEE Trans. Wireless Commun.*, 2015.

[146] Weimin Xiao et al. Uplink power control, interference coordination and resource allocation for 3GPP E-UTRA. In *Proc. of IEEE 64th Vehicular Technology Conference*, pages 1–5, Sep. 2006.

[147] Hong Yang and T.L. Marzetta. Performance of conjugate and zero-forcing beamforming in large-scale antenna systems. *IEEE J. Sel. Areas Commun.*, 31(2):172–179, Feb. 2013.

[148] Hong Yang and T.L. Marzetta. Total energy efficiency of cellular large scale antenna system multiple access mobile networks. In *In Proc. of IEEE Online Conference on Green Communications (GreenCom)*, pages 27–32, Oct. 2013.

[149] Haifan Yin, D. Gesbert, and L. Cottatellucci. Dealing with interference in distributed large-scale MIMO systems: A statistical approach. *IEEE Journal on Selected Topics in Signal Processing*, 8(5):942–953, Oct. 2014.

[150] Haifan Yin, D. Gesbert, M.C. Filippou, and Yingzhuang Liu. Decontaminating pilots in massive MIMO systems. In *Proc. of IEEE Int. Conf. Commun. (ICC)*, pages 3170–3175, June 2013.

[151] Dawei Ying, F.W. Vook, T.A. Thomas, D.J. Love, and A. Ghosh. Kronecker product correlation model and limited feedback codebook design in a 3d channel model. In *Proc. of IEEE ICC*, pages 5865–5870, June 2014.

[152] Su Khiong Yong and Chia-Chin Chong. An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges. *EURASIP J. Wireless Commun.*, 2007:1–10, 2006.

[153] Xianghao Yu, Juei-Chin Shen, Jun Zhang, and Khaled Letaief. Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems. 2016.

[154] Wence Zhang, Hong Ren, Cunhua Pan, Ming Chen, R.C. de Lamare, Bo Du, and Jianxin Dai. Large-scale antenna systems with UL/DL hardware mismatch: Achievable rates analysis and calibration. *IEEE Transactions on Communication*, 63(4):1216–1229, April 2015.

[155] X. Zhang, A.F. Molisch, and S.Y. Kung. Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection. *IEEE Transactions on Signal Processing*, 53(11):4091–4103, Nov. 2005.

[156] Kan Zheng, Suling Ou, and Xuefeng Yin. Massive MIMO channel models: A survey. *International Journal of Antennas and Propagation*, 2014:1–10, 2014.

Vita

Tianyang Bai received the Bachelor of Engineering in Communications Engineering from Harbin Institute of Technology, China in 2011, ranking first in the Honors School. He joined the Wireless Networking and Communication Group (WNCG) at The University of Texas at Austin (UT Austin) in 2011, and obtained his Masters of Science in Engineering in 2013. Since then, he has been pursuing his PhD in Electrical Engineering at UT Austin. He was an engineering research intern at the Huawei Technologies, Rolling Meadows, IL, in 2013, and at Qualcomm, Bridgewater, NJ, in 2014 and 2015. His research interests include applications of stochastic geometry, millimeter wave communications, and 5G cellular networks.

Email address: tybai@utexas.edu

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.